AD_____

Award Number: W81XWH-04-1-0082

TITLE: PARK2, a Large Common Fragile Site Gene, Is Part of a Stress Response Network in Normal Cells that is Disrupted during the Development of Ovarian Cancer

PRINCIPAL INVESTIGATOR: David I. Smith, Ph.D.

CONTRACTING ORGANIZATION: Mayo Foundation
Rochester, MN 55905

REPORT DATE: December 2005

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
01-12-2005	Annual	15 Dec 2004 – 14 Nov 2005

4. TITLE AND SUBTITLE

Development of PARK2, a Large Common Fragile Site Gene, Is Part of a Stress Response Network in Normal Cells that is Disrupted during the Ovarian Cancer

5a. CONTRACT NUMBER

5b. GRANT NUMBER
W81XWH-04-1-0082

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

David I. Smith, Ph.D.

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Mayo Foundation
Rochester, MN 55905

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES
Original contains colored plates: ALL DTIC reproductions will be in black and white.

14. ABSTRACT

PARK2 (Parkin) is a large gene spanning 1.3 megabases of genomic sequence within the unstable FRA6E (6q26) CFS. The two questions addressed here are what role the inactivation of Parkin plays in the development of ovarian cancer and whether this gene functions as part of a stress response network. First, we analyzed the effect of reintroducing Parkin into ovarian cancer cell lines lacking expression. Re-introduction of Parkin is associated with greater sensitivity to apoptotic induction. Utilizing real-time RT-PCR, we measured the expression of seven large genes residing within CFSs, including Parkin, in panels of cancer cell lines and primary tumors of the prostate, ovary, breast, brain and liver. This reveals non-random inactivation of these genes in cancers with greater inactivation of the large CFS genes in cancers that have a poorer clinical prognosis. This may offer a prognostic test of individual ovarian cancers based upon the number of large CFS genes inactivated in each cancer. Next, we examined Parkin as a stress response gene within cells. We utilized genome tiling arrays to characterize transcripts within and around Parkin and their response to two stresses. These studies reveal non-coding transcripts within large CFS genes. These studies support our overall hypothesis that the large CFS genes function as a stress response system within cells that are uniquely susceptible to genomic instability.

15. SUBJECT TERMS
Common fragile sties; Parkin; stress response

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			USAMRMC
U	U	U	UU	36	19b. TELEPHONE NUMBER *(include area code)*

Table of Contents

PARK2, a large common fragile site gene is part of a stress response network in normal cells that is disrupted during development of ovarian cancer

David I Smith, Ph.D. (P.I.)

Abstract

PARK2 (Parkin) is an extremely large gene that spans greater than 1.3 megabases of genomic sequence within chromosomal band 6q26. This gene is derived from within the middle of the highly unstable FRA6E common fragile site (CFS). CFSs are large chromosomal regions that are highly unstable and prone to deletions and other alterations, especially in developing cancer cells. The central two questions that we want to address with this work are what role does the inactivation of Parkin play in the development of ovarian cancer and whether this gene functions as part of a stress response network. In order to address these two questions, we have analyzed the effect of reintroducing Parkin into ovarian cancer cell lines that do not express it. We have already demonstrated that the re-introduction of Parkin is associated with greater sensitivity to the induction of apoptosis. This is consistent with our hypothesis that inactivation of this gene contributes to ovarian cancer development. We have now identified 20 extremely large genes like Parkin that reside within CFS regions. To determine if these genes are randomly inactivated during cancer development, we have utilized real-time RT-PCR to measure the expression of seven of these genes, including Parkin, in panels of cancer cell lines and primary tumors of the prostate, **ovary**, breast, brain and liver. This analysis reveals a decidedly non-random inactivation of the expression of these genes in different cancers. In addition, we've found that there is greater inactivation of expression of the large CFS genes in cancers that are more aggressive and have a poorer clinical prognosis. This may offer a prognostic test of individual ovarian cancers based upon the number of large CFS genes that are inactivated in each cancer. The second part of our studies was to examine Parkin as a stress response gene within cells. We have utilized the newly developed genome tiling arrays which contain tiling oligonucleotides across the non-redundant portion of genome to characterize transcripts within and around Parkin and their response to two stresses, hypoxia and treatment with the carcinogen NNK. These studies reveal that there are non-coding transcripts within large CFS genes and may begin to explain why these genes are so large in the first place. These studies support our overall hypothesis that the large CFS genes function as a stress response system within cells that are uniquely susceptible to genomic instability.

Introduction

Parkin is a gene that spans an extremely large chromosomal region of 1.36 Mbs. This large gene spans the most unstable region within the highly unstable FRA6E common fragile site (CFS) and our novel hypothesis that received Department of Defense Ovarian Cancer Research Program funding was that Parkin and other large CFS genes were part of a stress response system that is disrupted during the development of ovarian cancer. There were two major goals of this proposal. The first was to determine if inactivation of the expression of Parkin could contribute to the development of ovarian cancer. The second was to demonstrate that Parkin and other large common fragile site genes functioned as part of a stress response system within cells. We summarize the work that we've now finished in the second year of funding. Our key findings are the identification of an entire family of very large CFS genes similar to FHIT, WWOX and Parkin, and the demonstration that these genes are non-randomly inactivated in different cancers. We also identified the retinoic acid receptor-related orphan receptor alpha (RORA) as a large CFS gene whose expression is abrogated in many different cancers including ovarian cancer. This nuclear receptor transcription factor is extremely interesting because in addition to regulating many key cellular functions, it also appears to function as a stress regulated gene. We have also initiated a very

novel experiment to characterize Parkin and other large CFS genes as part of a stress response system within cells utilizing the newly developed genome tiling arrays.

Body

We again would like to thank the Department of Defense for their support of our work. We believe that the characterization of the common fragile sites (CFSs) is important because these large regions of genomic instability are highly sensitive, especially in developing ovarian cancers. In addition, many of these regions contain novel tumor suppressor genes which participate in ovarian cancer development. The first two genes identified within these unstable chromosomal regions were FHIT and WWOX. These genes have a highly unusual genomic organization as they both span extremely large genomic regions greater than 1.0 Mb. In spite of this, the final processed transcripts encoded by these genes are relatively small (1 Kb for FHIT, and 2.0 Kb for WWOX), thus the majority of these genes (greater than 99.8%) are intronic sequences. These two genes have been demonstrated to be tumor suppressor genes, both *in vitro* and *in vivo*. In addition, the inactivation of expression of these genes is associated with a poorer clinical outcome. Finally, both genes appear to be involved in cellular responses to stress.

The original goal of this proposal was to characterize another large CFS gene, the 1.36 Mb Parkin gene. This gene spans the third most unstable CFS region, FRA6E (6q26). Our two main goals in this project were: (A) to characterize the Parkin gene in ovarian cancers and to determine if the inactivation of Parkin was a frequent event in ovarian cancers and to determine if this had any functional significance; and (B) To determine if Parkin and other CFS genes were involved in the cellular responses to stress. We completed the first specific aim in the first year of this proposal where we demonstrated that the re-introduction of Parkin into an ovarian cancer cell line that did not express Parkin resulted in growth inhibition and also protected cells from mitochrondria-independent apotosis induced by ceramide. This work was published in *Oncogene* (1). In order to characterize Parkin as a gene involved in stress response, we have just initiated some novel studies utilizing the newly developed genome tiling arrays. These contain tiled oligonucleotides across the entire non-redundant portion of the genome and will enable us to interrogate not only the very small Parkin exons and their response to different stresses but the extremely large Parkin introns. Our hypothesis is that the large CFS genes, like FHIT, WWOX and Parkin, contain such large introns because they encode non-coding transcripts which respond to different stresses and regulate the expression of the CFS genes.

In this report we summarize our work where we have now identified an entire family of extremely large CFS genes. One CFS gene that we've now identified is RORA, the retinoic acid orphan receptor alpha. This 730 Kb gene spans the center of the FRA15A CFS (15q22.2) and is an extremely interesting nuclear transcription factor involved in the regulation of a number of key cellular processes. In addition, this gene is a cellular stress response gene. We describe our work characterizing the RORA gene where we demonstrate that this gene is frequently inactivated in multiple cancers including ovarian cancer and that it is involved in cellular stress response. The work on RORA was recently published in *Oncogene* (2).

When we originally wrote this proposal, we wanted to characterize how large genes like FHIT, WWOX, and Parkin could be responding to cellular stress. We could not imagine that a powerful technology like tiling arrays would be developed which in a single experiment would enable us to probe the entire genomes response to stress. We have been beta-testing the new 35 bp genome tiling arrays (which contain tiling oligonucleotides spaced 35 bp apart across the entire non-redundant portion of the human genome) and have devised an experiment to measure both coding and non-coding transcripts across the entire genome and their response to two types of stress, growth under hypoxic conditions and exposure

to the carcinogen NNK. We have completed this experiment and are beginning to analyze the huge amount of data generated. In the next year, we will be able to determine whether these stresses cause changes in non-coding transcripts which are present within the large introns of CFS genes like Parkin. Our hypothesis is that the reason there are such large genes within the highly unstable CFS regions is that the CFS regions are able to somehow transduce different cellular stresses into the production of the appropriate non-coding transcripts which then regulate the expression of the large CFS genes.

In this report, we therefore summarize our work on the identification of an entire family of large CFS genes, the identification and analysis of the RORA gene, and our preliminary studies utilizing tiling arrays to characterize the entire genomes response to stress.

Large genes within many CFS regions

Our analyses of several CFSs revealed that there were large genes (genes >1.0 Mbs) located within several, but not all, of these regions including FHIT (1.5 Mbs), Parkin (1.36 Mbs), GRID2 (1.39 Mbs), and WWOX (1.0 Mbs). In addition, along with others we demonstrated that FHIT, WWOX, and GRID2 were highly evolutionarily conserved and that the chromosomal regions surrounding them were also CFSs in mice (3-6). This suggested that the large gene and the unstable chromosomal region might be co-conserved because together they serve some function within cells.

We became interested in whether other CFS regions might also include large genes. To address this question, we collaborated with Dr. Robert Kuhn, a researcher at the UCSC Genome Database. Dr. Kuhn provided a list of all genes larger than 500Kb, and we carefully examined the list to remove redundant clones. We generated a list of 240 distinct human genes that spanned greater than 500 Kb of genomic sequence. These 240 genes represent the largest 1% of human genes.

Many of the largest human genes are derived from within CFS regions

Examination of the large gene list revealed that a number of these were derived from chromosomal bands that contained CFSs; we were curious how many corresponded to CFS genes. Our laboratory had already localized 20 of the 89 known CFS regions, and a few other CFS regions have been defined by other groups (7-10). A detailed examination of the sequences surrounding these localized CFSs identified several other large genes as CFS genes, including CNTNAP2 [this is the largest human gene which spans 2.3 Mb within 7q35 (FRA7I)] and LRP1B (1.9 Mb in FRA2F).

We then decided to test several of the largest human genes derived from chromosomal regions known to contain a CFS to determine if they were also CFS genes. A BAC clone covering the approximate center of each large gene was selected, labeled, and used as a FISH probe against metaphases prepared from cells exposed to aphidicolin. This analysis identified several other large CFS genes. However, not every large gene was derived from within a CFS region. Out of the 10 largest genes, 6 were determined to be derived from within CFS regions. Closer examination of the genomic region surrounding many of the localized CFS regions revealed that slightly less than half of the characterized CFS regions are associated with large genes. We can therefore estimate that there are approximately 40 large CFS genes distributed throughout the genome (11).

The Table on the following page lists the 20 known large CFS genes that have been identified as of today, the size of the genomic region spanned by each gene, the number of exons and the size of the final processed transcripts (FPT), the chromosomal location, and the CFS that spans each gene.

Gene Name	Size	Exons/FPT	Location	Fragile Site
CNTNAP2	2304258	25/8107	7q35	FRA7I
DMD	2092287	79/13957	Xp21.1	FRAXC
LRP1B	1900275	91/16556	2q22.1	FRA2F
CTNNA3	1775996	18/3024	10q21.3	FRA10D
DAB1	1548827	21/2683	1p32.3	FRA1B
FHIT	1499181	9/1095	3p17.2	FRA3B
KIAA 1680	1474315	11/5803	4q22.1	FRA4D
GRID2	1467842	16/3024	4q22.3	FRA4D
Dlg2	1463760	23/3071	11q14.1	FRA11F
Parkin	1379130	12/2960	6q26	FRA6G
IL1RAPL1	1368739	11/2722	Xp21.2	FRAXC
WWOX	1113013	9/2264	16q23.2	FRA16D
PDGFFA	917434	24/2550	4q12	FRA4B
IMMPL2	899238	6/1540	7q31.1	FRA7K
RORA	731967	11/1816	15q22.2	FRA15A
PTPR6	731390	30/4707	3p14.2	FRA3B
Neurobeachin	730417	58/10812	13q13.2	FRA13A
LARGE	647480	16/4326	22q12.3	FRA22B
ARHGAP15	638958	14/1706	2q22.2	FRA2F
SCAl	462345	9/10601	6p22.3	FRA6C

Similarities between the known large CFS genes

The large CFS genes share a number of similarities. Each of these genes is predominantly intronic (greater than 99.7%) and span some of the most unstable chromosomal regions in the genome which are difficult regions to transcribe as well as replicate. In addition, several of the large genes such as FHIT, WWOX, and GRID2 have been found to be highly conserved and the chromosomal regions surrounding them are fragile sites in mice. When comparing what little is known about the function of some of these large genes, it appears that many of them have completely different functions. However, one interesting connection shared by many of the large CFS genes is an association with normal neurological development.

This is already quite clear with Parkin, which when inactivated results in specific cellular death of cells that make and respond to dopamine leading to early onset juvenile Parkinson's disease (12-13). A spontaneous mouse mutant was identified that had a 1.0 Mb deletion within the distal portion of FRA6E. This deletion removed coding sequences of both Parkin and the immediately adjacent large Parkin co-regulated gene product PACRG. The mice that are homozygous for this deletion have a neurological phenotype known as *Quaking*(viable) (14), which have quake-like tremors caused by improper myelination of the CNS.

The GRID2 gene is a very large CFS gene which was first identified because of spontaneous deletions in mice resulting in a neurological defect known as *Lurcher* (3). Heterozygous *Lurcher* mice display ataxia as a result of selective, cell autonomous and apoptotic death of cerebellar Purkinje cells during postnatal development (15-16). This gene is also highly conserved between humans and mice and the region surrounding this gene is a CFS in the mouse (3), identical to what was observed for FHIT and WWOX.

Yet another large gene identified by us as a CFS gene is DAB1. Dab1 is the human homolog of the Drosophila disabled locus and it interacts with Reelin (17-18). When this locus is mutated in mice it results in a neurological defect known as *scrambler* (19). These mice have cerebellar hyperplasia with Purkinje cell ectopia. The normal function of Dab1 is to promote normal positioning of upper layer cortical plate neurons (20).

We tested a number of the largest human genes that were involved in neurological development, which were also derived from chromosomal regions known to contain a CFS and identified a number of other large CFS genes. This includes the Duchene Muscular Dystrophy (DMD) gene in FRAXC, LARGE in FRA22B (which is associated with myodystrophy when deleted in the mouse), and the Sca1 gene (associated with spinocerebellar ataxia) in FRA6C. Thus, many CFS genes are large genes that are involved in normal neurological development.

RORA and FRA15A

RORA is an orphan retinoic acid receptor and appears to be an important regulatory transcription factor involved in many pathophysiological processes such as cerebellar ataxia, inflammation, atherosclerosis, and angiogenesis. RORA is also the target for hypoxia-inducible factor 1, regulates plasma cholesterol levels, and positively regulates the expression of apolipoproteins A-I and C-III (21). RORA is a member of the steroid hormone nuclear receptor superfamily, which includes receptors for steroids, retinoids, and thyroid hormones (22). RORA was originally termed an orphan receptor because there was no knowledge about its natural ligands. However, subsequent studies have revealed various targets for RORA including fibrinogen-beta (23). RORA has also been shown to interact with NM23-2, a nucleoside diphoshate kinase involved in organogenesis and differentiation as well as NM23-1, the product of a tumor metastasis suppressor candidate gene (24).

RORA is the largest gene in the ROR family spanning over 730 Kb of genomic sequence within chromosomal band 15q22.2. RORA is composed of 11 small exons which together comprise a final processed transcript of 1816 bps; this is very similar in genomic organization to FHIT and WWOX. In addition, RORA is also linked to neurological development. The homozygous RORA mouse mutant, *staggerer*, has ataxia associated with cerebellar degeneration and a reduced number of Purkinje cells (25). *Staggerer* also displays other phenotypes such as dysfunction of smooth muscle cells and enhanced susceptibility to atherosclerosis (26).

Because RORA shares similarities with other well studied large CFS genes and is located in a chromosomal band (15q22.2) known to contain the FRA15A CFS, we tested whether RORA was in fact a CFS gene. This analysis revealed that RORA spanned the middle of the FRA15A region. Figure 1 below shows representative FISH results with a BAC from the middle of the RORA gene demonstrating that in one metaphase (A) the BAC hybridizes distal to the region of decondensation/breakage and in another metaphase (B) the BAC hybridizes proximal.

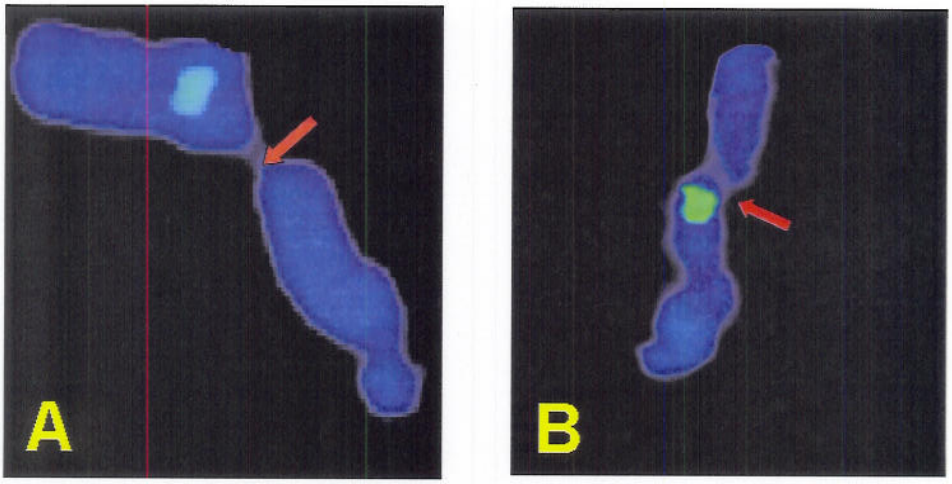


Figure 1 Depiction of FISH results obtained with a BAC clone crossing the middle of RORA and determined to be crossing FRA15A. BAC clone CTD-2034M3 was labeled with biotin and hybridized to normal human lymphocytes treated for 24 hours with 0.4μM aphidicolin. 20 metaphases with clear breakage/decondensation at 15q22.2 were scored. The hybridization signal appeared proximal to the break in 12 metaphases and distal in 8, showing that RORA is located in the approximate center of FRA15A. A. Representative metaphase with the hybridization signal appearing distal to the break. B. Representative metaphase with the hybridization signal appearing proximal to the break.

Out of 20 metaphases with good discernible breakage within 15q22.2, we found that this BAC hybridized proximal to the region of breakage 12 times and distal 8 times. This finding would place this BAC clone and the RORA gene itself within the middle and most unstable region of FRA15A.

According to one previously published study, there are four RORA isoforms (RORA 1, 2, 3, and 4) which are produced by alternative splicing (27). In our studies in various normal tissues, we found expression of only RORA 1 and 4. Figure 2 below shows the transcriptional level of RORA 1 and 4 in various normal tissues.

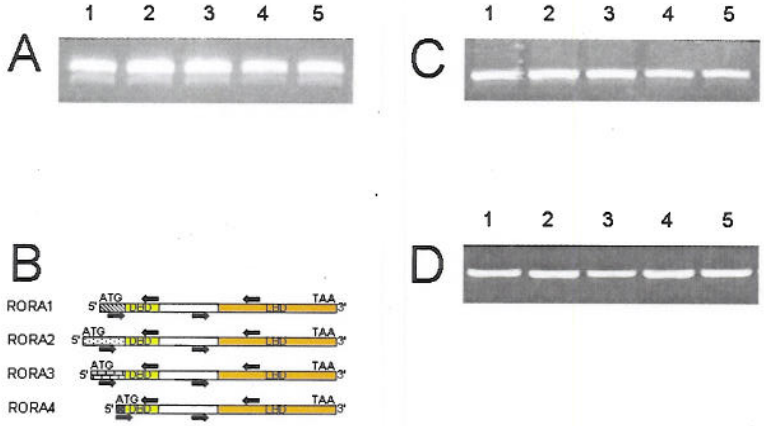


Figure 2 The transcriptional level of RORA and its isoforms in various normal tissues. Lane 1 brain; Lane 2 breast; Lane 3 liver; Lane 4 ovary; Lane 5 prostate. Total RNA was prepared from normal human tissues and cDNA was generated. Semi-quantitative RT-PCR was performed using the universal primers for all RORA isoforms and the specific primers for each isoform to measure the level of RORA. A. RORA universal primers; B. A schematic diagram showing the four different isoforms of RORA. The bold arrows show the positions of specific primers for each isoform and the universal primers for all isoforms. DBD DNA binding domain; LBD ligand binding domain. C. RORA1 (isoform 1) primers; D. RORA4 (isoform 4) primers.

We next examined the expression of RORA in several different types of human cancer samples, either in primary tumors or in tumor-derived cell lines using RT-PCR. This revealed that RORA was down-regulated in breast, prostate, and **ovarian** cancers (see Figure 3 below). These results are consistent with those obtained in studies of other critical CFS genes including FHIT (28-31), WWOX (32-34), and Parkin (1).

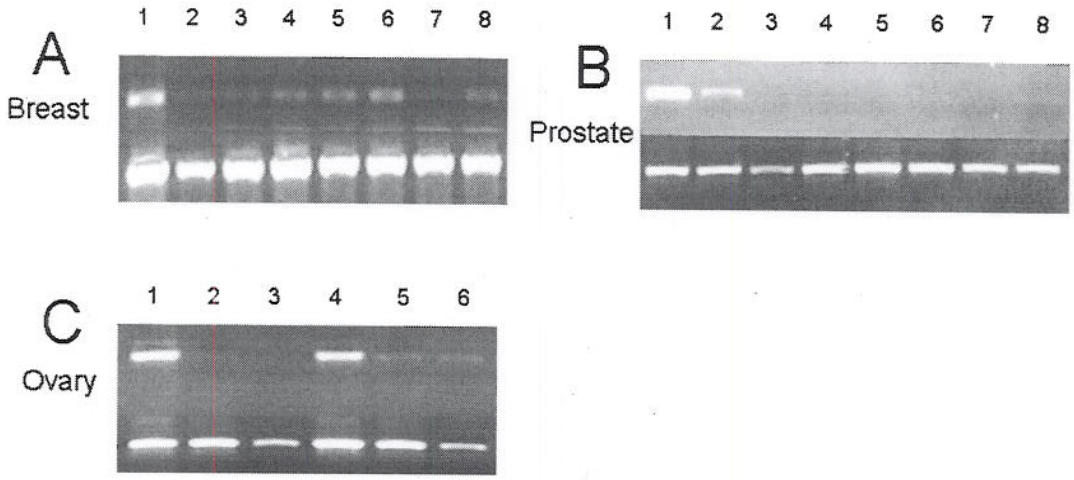


Figure 3. RORA is down-regulated in different types of cancer cell lines and human primary cancers. Total RNA was extracted and reverse transcribed into cDNA. PCR was performed with RORA universal primers to examine the transcriptional level of RORA. A. Breast cancer cell lines, Lane 1 MCF12F; Lane 2 MCF7; Lane 3 MDA157; Lane 4 UACC893; Lane 5 ZR75; Lane 6 MDA435; Lane 7 T47D; Lane 8 BT474. Top row RORA; Bottom row Actin. B. Prostate cancer cell lines and primary tumor samples, Lane 1 normal prostate control; Lane 2 DU145; Lane 3 PC3; Lane 4 LNCaP; Lane 5-8 primary prostate tumor tissues. Top row RORA; Bottom row Actin. C. Ovary cancer cell lines, Lane 1 normal ovarian epithelium control (OSE); Lane 2 OV167; Lane 3 OV177; Lane 4 OV202; Lane 5 OVCAR5; Lane 6 SKOV3. Top row RORA; Bottom row Actin.

We also examined whether RORA expression was modulated by different types of cellular stress other than hypoxia. We first demonstrated that RORA is activated by exposure to aphidicolin (Figure 4) and then subsequently showed similar activation by other types of cellular stress including exposure to ultraviolet radiation (UV), addition of the carcinogen MMS (methyl-methane sulfonate), and treatment with H_2O_2 (oxidative stress) (2). Figures 4 and 5 below display the changes in RORA transcripts and RORA protein levels in response to some of these stresses.

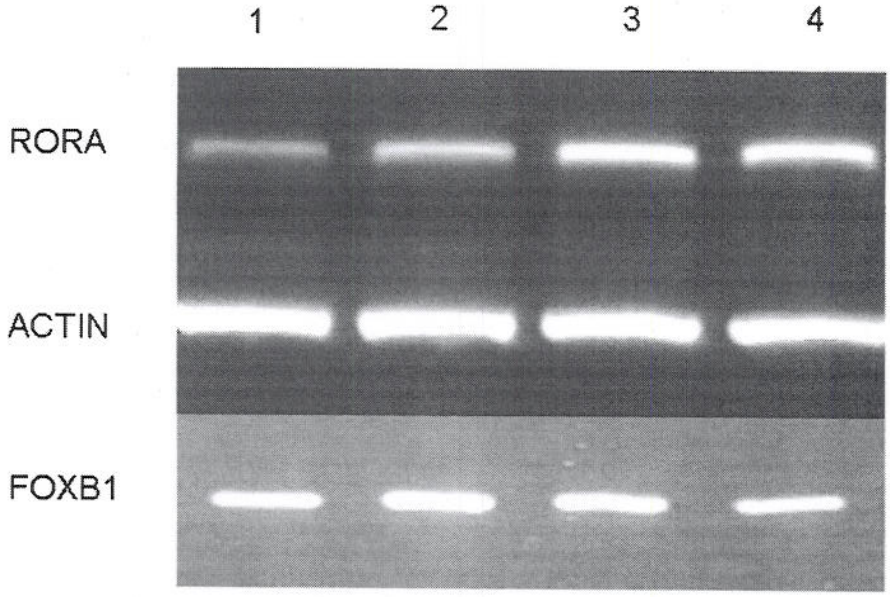


Figure 4 The effect of aphidicolin (APC) on the transcription of RORA in MCF12F cells. MCF12F cells were treated with various doses of APC for 24 hours before total RNA was extracted and cDNA was prepared. PCR was set up to check the transcriptional level of RORA, FOXB1 (a gene within FRA15A right next to RORA) and Actin, using the universal primers for RORA, primers for FOXB1 and the control primers for Actin. Lane 1 cell without APC treatment; Lane 2 with APC 0.2 μM; Lane 3 with APC 0.4 μM; Lane 4 with APC 0.8 μM.

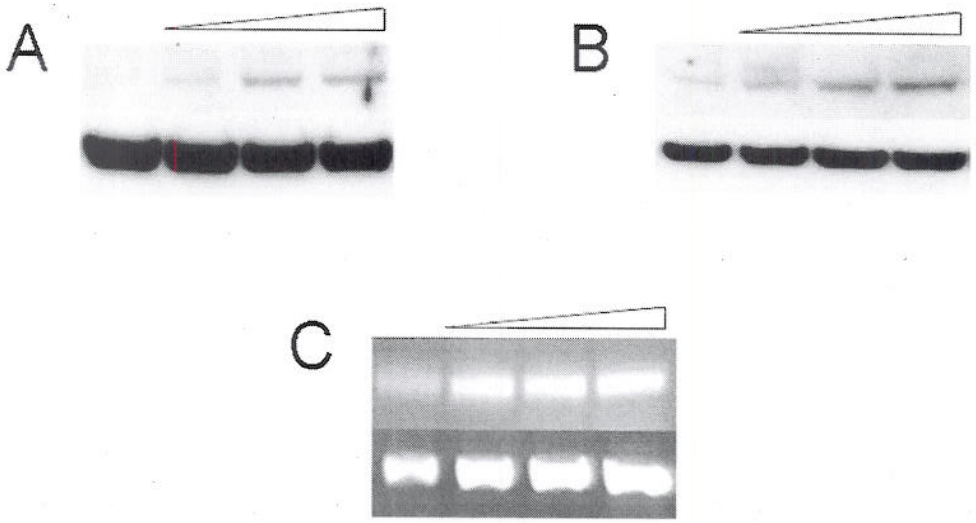


Figure 5. The expression of RORA in MCF12F cells is activated by different types of stress treatments. A. The effect of UV on the protein level of RORA. MCF12F cells were treated with UV at 10, 20 and 50 J/m^2 and total protein was prepared. B. The effect of MMS on the protein level of RORA. MCF12F cells were treated with MMS at 0.001%, 0.005% and 0.01% for 24 hours. The level of Rora was examined with anti-Rora antibody. C. The effect of H_2O_2 on the transcriptional level of RORA. MCF12F cells were treated with H_2O_2 at 100, 200 and 500 μM for 24 hours and then total RNA was extracted and cDNA was prepared. RT-PCR was performed using universal primers for RORA.

An important question is what role alterations in expression of this large CFS gene play in the development of breast cancer? Indeed all of the CFS genes could be frequent targets of alterations in unstable cancer cells because of the unstable regions that surround them. We transfected RORA into the breast cell line MCF12F and found that increased RORA expression resulted in decreased growth of MCF12F cells (see Figure 6 below). These results are similar to those obtained with FHIT, WWOX, and Parkin.

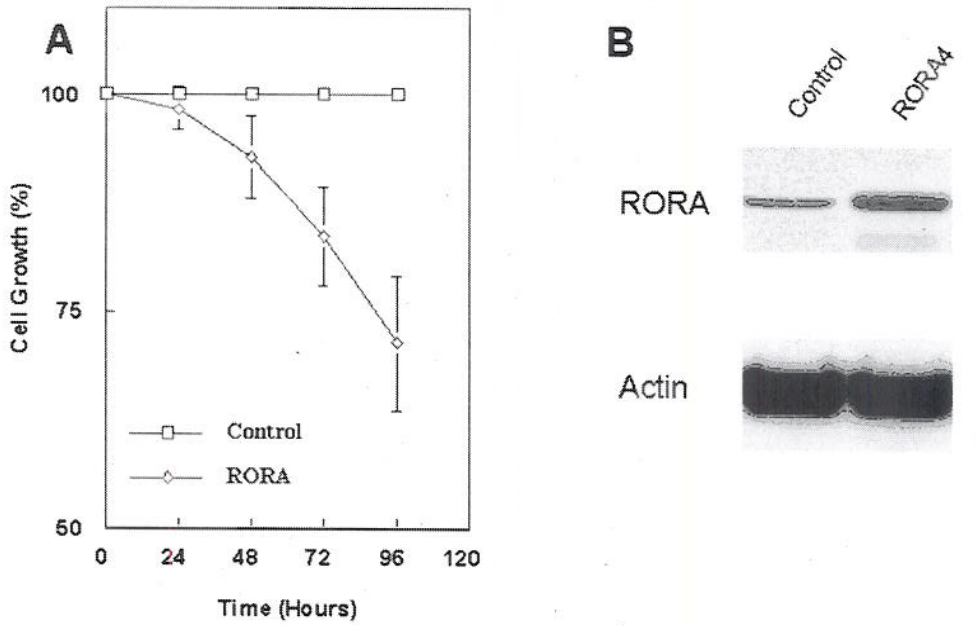


Figure 6 The effect of RORA over-expression on cellular growth. A. MCF12F cells were plated and incubated overnight before transfection. The plasmids (pcDNA3 as control and pcDNA3-RORA4) were transfected into cells using a Lipofectamine 2000 transfection kit following the manufacturer's protocol. The cell number was counted 24, 48, 72 and 96 hours later. All results are the average of at least three independent experiments with standard deviations shown by bars. B. The level of RORA in pcDNA3-RORA4 transfectants detected by the Western blotting assay. Top row RORA; Bottom row Actin.

Thus, changes in RORA expression are associated with readily observable changes in growth rates of MCF12F cells, which supports our contention that inactivation of RORA expression could provide a significant growth advantage to cells thus participating in breast cancer development. All of this work is summarized in our recent paper that was published in *Oncogene* (2).

Expression of the large CFS genes in cancers and cancer-derived cell lines

Our hypothesis is that the large CFS genes are part of a stress response system within cells that is uniquely susceptible to genomic instability and that in cancers with considerable genomic instability, there will be inactivation (alterations) of expression of multiple CFS genes. We have already demonstrated observable phenotypic changes associated with alterations in the expression of these genes. Next we sought to determine whether these genes were randomly inactivated or whether there might be some selection for inactivation of specific CFS genes in different cancers.

To address this question, we used real-time RT-PCR to precisely measure the expression of seven representative large CFS genes (FHIT, WWOX, **Parkin**, GRID2, DLG2, DAB1 and the two expressed RORA isoforms 1 and 4) in panels of primary tumors and cancer-derived cell lines for cancers of the prostate, breast, ovary, liver, and brain. PCR primers were constructed to be optimal for real-time RT-PCR analysis (100-125 bp products derived from the 3' end of the final processed transcripts from these genes) and then we performed real-time RT-PCR in the ABI 7900 real-time PCR machine. To quantify the expression of each of these genes (we constructed primers to differentiate between the two RORA isoforms), we compared the C_t measurements obtained with each gene to that of the β–actin gene and used the delta C_t measurements to quantify message amounts for the large CFS genes. We obtained several normal tissues for comparison for each tissue/tumor type and compared the expression of β-actin and the CFS genes in those normal tissues to panels of cancer-derived cell line, as well as primary tumors of that same type. We considered any gene to be aberrantly regulated if its expression was more than 4-fold up or down relative to the range of expression determined for the normal samples after each sample was run in triplicate.

We found that the expression of these genes was frequently abrogated in different cancers and there appeared to be a very non-random pattern of gene inactivation. We also observed that many cancers had inactivation of multiple large CFS genes. Those cancers with a great deal of genomic instability will have inactivation of many of these genes simultaneously which could have a profound phenotypic effect on those cells. For each of the CFS genes tested, the Table below indicates the number of primary tumors/cell lines that had decreased expression compared to normal samples divided by the total number of tumors/cell lines tested.

	FHIT	WWOX	**Parkin**	Grid2	Dlg2	Dab1	RORA1	RORA4
Prostate	0/17	1/17	1/17	0/17	2/17	0/17	10/17	1/17
Ovary	**2/18**	**1/18**	**3/18**	**1/18**	**12/18**	**0/18**	**2/18**	**5/18**
Breast	4/16	3/16	5/16	0/16	8/16	3/16	3/16	**8/16**
Brain	7/17	10/17	5/17	9/17	17/18	11/17	5/17	0/17
Liver	4/15	11/15	14/15	14/15	12/15	12/15	15/15	0/15

It is important to note that there was a very interesting preliminary correlation between the frequency of inactivation of these CFS genes and cancers that have very poor clinical outcomes. We found the least inactivation of CFS gene expression in cancers of the prostate, which of the various cancers examined has the best clinical outcome. There was much greater loss of expression of these genes in cancers of the breast and ovary, and many of these cancers tend to be more aggressive than prostate cancers. However,

the cancers with the greatest inactivation of these genes were cancers of the brain and liver. These cancers are highly aggressive, and there is a high probability that patients who develop these tumors will succumb to them.

Monitoring the entire genomes response to stress using Tiling Arrays

Much of the focus in cancer genetics has been the identification of important alterations during cancer development which could contribute to that process. This focus has been primarily upon the genes and more importantly upon the coding portions of those genes. However, only 5% of the genome corresponds to the genes themselves and less than 1% of the genome itself corresponds to the exons or coding portions of those genes. Nowhere is this discrepancy between genome size and apparent coding potential more evident than in some of the large CFS genes. The 1.36 Mb Parkin gene produces a 2960 bp final processed transcript, making this gene over 99.8% intronic. Why produce such a large initial transcript only to process it down to such a small final processed transcript? One possibility is that there may be regulatory RNAs produced from the intronic sequences which regulate the expression of Parkin (similar to the prostate susceptibility locus within one of the large FHIT introns). However, with no knowledge of where within the large Parkin introns such transcripts are derived, there is no feasible way to identify these potential stress-regulated transcripts.

We were considering the construction of sufficient oligonucleotides to probe across the large Parkin introns in order to identify putative non-coding transcripts at a cost of many thousands of dollars when Affymetrix began beta-testing their new tiling arrays which contain tiled oligonucleotides across the entire non-redundant portion of the genome (The Human Tiling 1.0R Array Set). They produce a 5 bp tiling array which contains overlapping oligos across the genome whose centers are 5 bp apart and 35 bp tiling arrays (where the 25-mers centers are 35-bp apart). The current version of microarray chips have 6.5 million features; it was possible for Affymetrix to completely cover the non-redundant portion of the entire human genome with the 35 bp tiling arrays on 14 chips (these chips also have both perfect match and mismatch oligos at each position, similar to what is present on the U133 Plus2 arrays for gene expression analysis). We obtained these chips as part of the β-testing of the 35 bp tiling array and realized that we could not only examine the entire 1.36 Mb region containing Parkin, but we can now monitor the entire genomes response to stress.

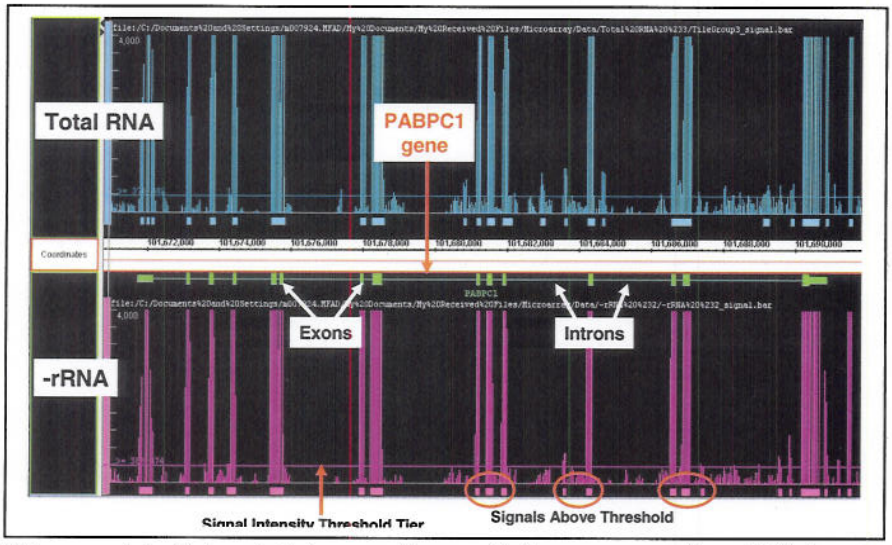
The tiling arrays have been pioneered by Dr. Tom Gingeras and co-workers at Affymetrix. Using the 5 bp tiling arrays for 10 human chromosomes, they demonstrated that unannotated, nonpolyadenylated transcripts comprise the major proportion of the transcriptional output of the human genome (35). This provides additional support for our hypothesis that the non-coding portion of the genome may still be transcriptionally active and thus the large introns may produce important transcripts within the large CFS genes like Parkin.

Our microarray experiment was set up to measure the entire genomes response to two different types of cellular stress, growth under hypoxic conditions, and exposure to the carcinogen from cigarette smoke 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone (NNK) (36). We cultured normal ovarian surface epithelial cells and exposed them to hypoxia or NNK. Hypoxia is a physiologically important endoplasmic reticulum (ER) stress that is present in all solid tumors (37). Hypoxia can influence tumor cells in one of two ways, either by acting as a stressor that impairs growth or causes cell death (slowing of proliferation, apoptosis, or necrosis) (38) or by serving as a factor that ultimately results in malignant progression and increased resistance to radiation therapy and other cancer treatments (39). In contrast, NNK is a compound formed by the nitrosation of nicotine and has been identified as the most potent carcinogen in cigarette smoke (36). Once NNK is bioactivated by cytochrome P450 metabolizing

enzymes (CYP2A6 and 2A13) (40), it can induce DNA damage, form DNA adducts (41), increase oxidative stress (42), as well as induce p53 and RAS mutations (43).

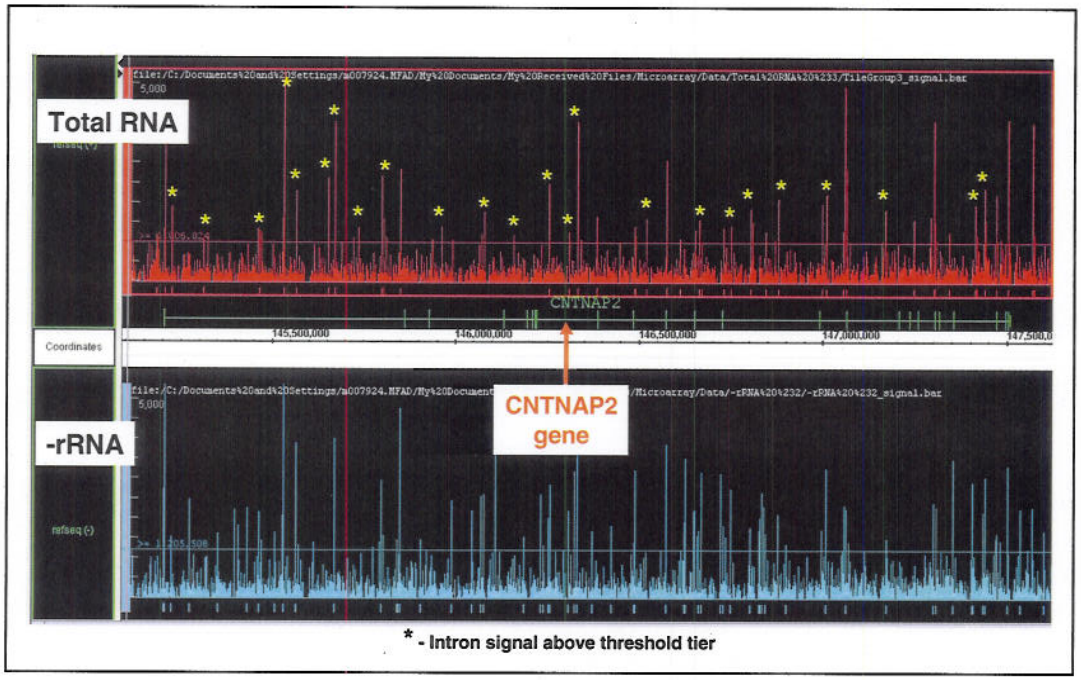
This work was done in collaboration with the Microarray Core of the Mayo Clinic and together, we went through the protocol for producing cDNA from all the RNA species present using random oligonucleotide primers (since we want to label more than just the polyA+ mRNAs). One potential problem with this experiment is that total RNA contains a vast excess of ribosomal RNA which could potentially swamp out signals coming from intronic regions. There is a commercially available kit using magnetic beads which can purify away the ribosomal RNA (the RiboMinus kit), and we compared the hybridization of cDNA produced from total RNA to cDNA produced from RiboMinus purified RNA. The RiboMinus protocol resulted in enrichment of non-ribosomal RNA but also resulted in some degradation of the remaining RNA species. We hybridized equivalent amounts of labeled total RNA and RiboMinus purified RNA (-rRNA) to tiling array chips so that we could compare the hybridization signals to help determine whether the additional expense of the RiboMinus purification kits was worthwhile. We took a specific chip that contained oligos across portions of chromosomes 7 and 8 (as we had extra copies of that chip from Affymetrix) and could then examine the signals obtained when that chip was hybridized with labeled total RNA as compared to labeled -rRNA. There are a number of control genes present on the chip so that we could compare the intensity of hybridization signals between the two different labeled RNAs.

The Figure below shows a normalization control: the 19.2 Kb gene for poly(A) binding protein PABPC1. The PABPC1 gene itself is represented by the horizontal green bar in the middle of the diagram (bold green indicating exons and the intron regions are displayed by the thin green line). The top panel shows the hybridization with total RNA (in red) and the bottom the hybridization with –rRNA (cyan). The bold colored bars under the bar graph for both samples correspond to the regions of PABPC1 where the signal intensity is considered significant. The Figure demonstrates that the hybridization signals are quite comparable between the two RNA samples. There are some signals coming from the introns of this gene, but the signal intensity of most of them is too low and is below the cut-off bar seen on both graphs labeled Signal Intensity Threshold Tier.



We ran chip 7 (containing portions of chromosome 7 and 8) because we had two extra copies of this particular chip which could be used to compare the hybridization of the two different labeled RNAs. Chromosome 7 also contains the largest human gene (the 2.3 Mb CNTNAP2 gene) which is also a CFS gene. The Figure below shows the hybridizations obtained with the two labeled RNAs across this gene. Here, too, there are very similar hybridizations between the two labeled RNAs and the majority of the

strong signals are coming from the small exons of this gene. However, it is interesting to note that there are a number of signals observed within the large intronic regions of this gene. Signals above threshold are shown with a star with the hybridization done with total RNA. The 25 exons of the CNTNAP2 gene are shown as green bars in the top figure with hybridization to total RNA. All 25 exons produce significant signals but in addition, there are a number of strong signals that are clearly coming from intronic sequences within this large gene. These are the potential signals that we are going to be interested in coming from the large introns of the RORA gene.



Based upon the results obtained with this preliminary chip hybridization, we decided to forgo the use of the RiboMinus kits and simply label total RNA for the tiling arrays. The labeling and hybridizations are ongoing as this proposal is being put together. We have designed the experiment with consultation from our statistical colleagues. Details concerning the experimental design are discussed further in the *Materials and Methods* section.

Our goal will be to examine the entire 1.36 Mb Parkin gene and then to determine if any of the non-coding transcripts produced from within the large introns of this gene have changes in expression in response to either type of cellular stress. We would then characterize these transcripts in greater detail and also determine if there are alterations in these transcripts in panels of primary ovarian cancers. One of the best things about the tiling array experiment is that it enables us to examine the entire genome's response to stress. In the future we could expand our analysis to determine first how each of the large CFS genes responds to stress and then how the remainder of the genome responds to stress.

Key Research Accomplishments

We have made several major research accomplishments in the past year of this grant. The first was the demonstration that there is actually a family of extremely large CFS genes. We have now identified a total of 20 large CFS genes similar to FHIT, WWOX, and Parkin. We have identified the 730 Kb RORA gene as one of the CFS genes. This important nuclear transcription factor is involved in the regulation of a number of key cellular processes, and we have shown that the expression of this gene is frequently inactivated in many ovarian cancers. One important concern is whether the large CFS genes are inactivated in different cancers simply because they reside within the highly unstable CFS regions and

are therefore passengers to the instability of the regions that surround them. We have now demonstrated that these genes are non-randomly inactivated in different cancers. We have also shown that while there is infrequent inactivation of expression of these genes in prostate cancers, which generally have a good clinical prognosis, these genes are inactivated much more frequently in ovarian cancers. However, in cancers of the liver and brain which have very poor clinical prognoses, we find that many of the CFS genes are inactivated simultaneously. This demonstrates that these genes are not inactivated simply because of the unstable regions that they reside in and that there appears to be a selection for inactivation of these genes. It also suggests that an analysis of ovarian cancers for inactivation of these genes may be a useful diagnostic for individual cases of ovarian cancers. We would expect that those cancers that have inactivation of multiple large CFS genes would have a poorer clinical outcome. We propose to examine this in this final year of this grant.

We are also utilizing the very powerful technology of genome tiling arrays to characterize the total genomes response to two different types of stress, growth under hypoxic conditions and exposure to the carcinogen NNK. This experiment has been completed, and we are just beginning to analyze the huge amount of data generated by this experiment to characterize both coding and non-coding transcripts that respond to stress. We anticipate that we will be able to identify important non-coding transcripts from within the large Parkin gene that respond to stress. This would support our overall hypothesis that the function of the highly unstable CFS regions and the large genes contained within them is as a genome-wide stress response system. This would explain why these very large genes within the most unstable chromosomal regions in the genome are so highly evolutionarily conserved.

Appendix

Smith DI, Zhu Y, McAvoy S, Kuhn R. Common fragile sites, extremely large genes, neural development and cancer. *Cancer Lett* 2006, 232: 48-57.

Zhu Y, McAvoy S, Kuhn R, **Smith DI**. RORA, a large common fragile site gene, is involved in cellular stress response. *Oncogene* 2006, In Press.

References

1) Denison SR, Wang F, Becker NA, Schule B, Kock N, Phillips LA, Klein C, Smith DI. Alterations in the common fragile site gene Parkin in ovarian and other cancers. *Oncogene* 2003, 22: 8370-8378.
2) Zhu Y, McAvoy S, Kuhn R, **Smith DI**. RORA, a large common fragile site gene, is involved in cellular stress response. *Oncogene*, 2006, In Press.
3) Rozier L, El-Achkar E, Apiou F, Debatisse M. Characterization of a conserved aphidicolin-sensitive common fragile site at human 4q22 and mouse 6C1: possible association with an inherited disease and cancer. *Oncogene* 23: 6872-6880.
4) Krummel KA, Denison SR, Calhoun E, Phillips LA, Smith DI. The common fragile site FRA16D and its associated gene WWOX are highly conserved in the mouse at Fra8E1. *Genes, Chromosomes Cancer*2002, 34: 154-167.
5) Glover TW, Hoge AW, Miller DE, Ascara-Wilke JE, Adam AN, Dagenais SL, Wilke CM, Dierick HA, Beer DG. The murine Fhit gene is highly similar to its human orthologue and maps to a common fragile site region. *Cancer Res* 1998, 58: 3409-3414.
6) Matsuyama A, Shiraishi T, Trapaso F, Kuroki T, Alder H, Mori M, Huebner K, Croce CM. Fragile site orthologs FHIT/FRA3B and Fhit/Fra14A2: evolutionarily conserved but highly recombinogenic. *Proc Natl Acad Sci USA* 2003, 100: 14988-14993.
7) Mishmar D, Rahat A, Scherer SW, Nyakatura G, Hinzmann B, Kohwi Y, Mandel-Gutfroind Y, Lee JR, Drescher B, Sas DE, Margalit H, Platzer M, Weiss A, Tsui LC, Rosenthal A, Kerem B. Molecular characterization of a common fragile site (FRA7H) on human chromosome 7 by the cloning of a simian virus 40 integration site. *Proc Natl Acad Sci USA* 1998, 95: 8141-8146.

8) Limongi MZ, Pelliccia F, Rocchi A. Characterization of the human common fragile site FRA2G. *Genomics* 2003, 81: 93-97.
9) Arlt MF, Miller DE, Beer DG, Glover TW. Molecular characterization of FRAXB and comparative common fragile site instability in cancer cells. *Genes, Chromosomes Cancer* 2002, 33: 82-92.
10) Morelli C, Karayianni E, Magnanini C, Mungall AJ, Thorland E, Negrini M, Smith DI, Barbanti-Brodano G. Cloning and characterization of the common fragile site FRA6F harboring a replicative senescence gene
11) Smith DI, Zhu Y, McAvoy S, Kuhn R. Common fragile sites, extremely large genes, neural development and cancer. *Cancer Lett*2006, 232: 48-57.
12) Kitada T, Asakawa S, Hattori N, Matsumine H, Yamamura Y, Minoshima S, Yokochi M, Mizuno Y, Shimizu N. Mutations in the parkin gene cause autosomal recessive juvenile parkinsonism. *Nature* 1998, 392: 605-608.
13) Hattori N, Kitada T, Matsumine H, Asakawa S, Yamamura Y, Yoshino H, Kobayashi T, Yokochi M, Wang M, Yoritaka A, Kondo T, Kuzuhara S, Nakamaura S, Shimizu N, Mizuno Y. Molecular genetic analysis of a novel Parkin gene in Japanese families with autosomal recessive juvenilie parkinsonism: evidence for variable homozygous deletions in the Parkin gene in affected individuals. *Ann Neurol* 1998, 44: 935-941.
14) Lorenzetti D, Bishop CE, Justice MJ. Deletion of the Parkin coregulated gene causes male sterility in the quaking(viable) mouse mutant. *Proc Natl Acad Sci USA* 2004, 101: 8402-8407.
15) Doughty ML, De Jager PL, Korsmeyer SJ, Heintz N. Neurodegeneraton in Lurcher mice occurs via multiple cell death pathways. *J Neurosci* 2000, 20: 3687-3694.
16) Selimi F, Lohof AM, Heitz S, Laloutte A, Jarvis CI, Bailly Y, Mariani J. Lurcher GRID2-induced death and depolarization can be dissociated in cerebellar Purkinje cells. *Neuron* 2003, 37: 813-819.
17) Howell BW, Gertler FB, Cooper JA, Mouse disabled (mDab1): a Src binding protein implicatd in neuronal development. *EMBO* 1997, 16: 121-132.
18) Rice DS, Sheldon M, D'Arcangelo G, Nakajima K, Goldowitz D, Curran T. Disabled-1 acts downstream of Reelin in a signaling pathway that controls laminar organization in the mammalian brain. *Development*1998, 125: 3719-3729.
19) Ware ML, Fox JW, Gonzalez JL, Davis NM, Lambert de Rouvroit C, Russo CJ, Chua SC Jr., Goffinet AM, Walsh CA. Aberrant splicing of a mouse disabled homolog, mdab1, in the scrambler mouse. *Neuron* 1997, 19: 239-249.
20) Deguchi K, Inoue K, Avila WE, Lopez-terrada D, Antalffy BA, Quattrocchi CC, Sheldon M, Mikoshiba K, D'Arcangelo G, Armstrong DL. Reelin and disabled-1 expression in developing and mature human cortical neurons. *J Neuropathol Exp Neurol* 2003, 62: 676-684.
21) Jetten AM. Recent advances in the mechanisms of action and physiological functions of the retinoid-related orphan receptors (RORs) *Curr Drug Targets Inflamm Allergy* 2004, 3: 395-412.
22) Evans RM. The steroid and thyroid hormone superfamily. *Science* 1988, 240: 889-995.
23) Chauvet C, Bois-Joyeux B, Fontaine C, Gervois P, Bernard MA, Staels B, Danan JL. The gene encoding fibrinogen-beta is a target for retinoic acid receptor-related orphan receptor alpha. *Mol Endocrinol* 2005, 19: 2517-2526.
24) Paravicini G, Steinmayr M, Andre E, Becker-Andre M. The metastasis suppressor candidate nucleotide diphosphate kinase NM23 specifically interacts with members of the ROR/RZR nuclear orphan receptor subfamily. *Biochem Biophys Res Comm* 1996, 227: 82-87.
25) Dussault I, Fawcett D, Matthyssen A, Bader JA, Giguere V. Orphan nuclear receptor RORA alpha-deficient mice display the cerebellar defects of staggerer. *Mech Dev* 1998, 70: 147-153.
26) Jarvis CI, Staels B, Brugg B, Lemaigre-Dubreuil Y, Tedgui A, Mariani J. Age-related phenotypes in the staggerer mouse expand the RORalpha nuclear receptor's role beyond the cerebellum. *Mol Cell Endocrinol* 2002, 186: 1-5.
27) Giguere V, Tini M, Flock G, Ong E, Evans RM, Otulakowski G. Isoform-specific amino-terminal domains dictate DNA-binding properties of ROR alpha, a novel family of orphan nuclear receptors. *Genes Dev* 1994, 8: 538-553.
28) Otterson GA, Xiao GH, Geradts J, Jin F, Chen WD, Niklinksa W, Kaye FJ, Yeung RS. Protein expression and functional analysis of the FHIT gene in human tumor cells. *J Natl Cancer inst*1998, 90: 426-432.
29) Sozzi G, Tornielli S, Tagliabue E, Sard L, Pezzella F, Pastorino U, Minoleeti F, Pilotti S, Ratcliffe C, Veronese ML, Goldstraw P, Huebner K, Croce CM, Pierotti MA. Absence of Fhit protein in primary lung tumors and cell lines with FHIT gene abnormalities. *Cancer Res* 57: 5207-5212.

30) Ingvarsson S. FHIT alterations in breast cancer. *Semin Cancer Biol* 2001, 11: 361-366.
31) Ingvarsson S, Agnarsson BA, Sigbjomsdottir BI, Kononen J, Kallioniemi OP, Barkardottir RB, Kovatich AJ, Schwarting R, Hauck WW, Huebner K, McCue PA. Reduced Fhit expression in sporadic and BRCA2-linked breast carcinomas. *Cancer Res* 1999, 59: 2682-2689.
32) Paige AJ, Taylor KJ, Taylor C, Hillier SG, Farrington S, Scott D, Porteous DJ, Smyth JF, Gabra H, Watson JE. WWOX, a candidate tumor suppressor gene involved in multiple tumor types. *Proc Natl Acad Sci USA* 98: 11417-11422.
33) Yendamuri S, Kuroki T, Trapasso F, Henry AC, Dumon KR, Huebner K, Williams NN, Kaiser LR, Croce CM. WW domain containing oxidoreductase gene expression is altered in non-small cell lung cancer. *Cancer Res* 2003, 63: 878-891.
34) Nunez MI, Ludes-Meyers J, Abba MC, Kil H, Abbey NW, Page RE, Sahin A, Klein-Szanto AJ, Aldaz CM. Frequent loss of WWOX expression in breast cancer: correlation with estrogen receptor status. *Breast Cancer Res Treat* 2005, 89: 99-105.
35) Cheng J, Kapranov P, Drenkow J, Dike S, Brubaker S, Patel S, Long J, Stern D, Tammana H, Helt G, Sementchenko V, Piccolboni A, Bekiranov S, Bailey DK, Ganesh M, Ghosh S, Bell I, Gerhard DS, Gingeras TR. Transcriptional maps of 10 human chromosomes at 5-nucleotide resolution. *Science* 2005, 308: 1149-1154.
36) Schuller HM. Mechanisms of smoking-related lung and pancreatic adenocarcinoma development. *Nat Rev Cancer* 2002, 2: 455-463.
37) Feldman DE, Chauhan A, Koong AC. The unfolded protein response: a novel component of the hypoxic stress response in tumors. *Mol Cancer Res* 3: 597-605.
38) Vaupel P, Harrison L. Tumor hypoxia: causation factors, compensatory mechanisms, and cellular response. *Oncologist* 2004, 9: 4-9.
39) Vaupel P, Kelleher DK, Hockel M. Oxygen status of malignant tumors: pathogenesis of hypoxia and significance for tumor therapy. *Semin Oncol* 2001, 28: 29-35.
40) Wang H, Tan W, Hao B, Miao X, Zhou G, He F, Lin D. Substantial reduction in risk of lung adenocarcinoma associated with genetic polymorphism in CYP2A13, the most active cytochrome P450 for the metabolic activation of tobacco-specific carcinogen NNK. *Cancer Res* 2003, 63: 8057-8061.
41) Cloutier JF, Drouin R, Weinfeld M, O'Conner TR, Castonguay A. Characterization and mapping of DNA damage induced by reactive metabolites of 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone (NNK) at nucleotide resolution in human genomic DNA. *J Mol Biol* 2001, 313: 539-557.
42) Bhagwat SV, Vijayasarathy C, Raza H, Mullick J, Avadhani NG. Preferential effects of nicotine and 4-(N-methlyl-N-nitrosamine)-1-(3-pyridyl)-1-butanone on mitochondrial glutathione S-transferse A4-4 induction and increased oxidative stress in the rat brain. *Biochem Pharmacol* 1998, 56: 831-839.
43) Cloutier JF, Drouin R, Weinfeld M, O'Conner TR, Castonguay A. Characterization and mapping of DNA damage induced by reactive metabolites of 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone (NNK) at nucleotide resolution in human genomic DNA. *J Mol Biol* 2001, 313: 539-557.

Smith, D.I.
Appendix 1

Available online at www.sciencedirect.com
SCIENCE DIRECT
ELSEVIER
Cancer Letters 232 (2006) 48–57
CANCER Letters
www.elsevier.com/locate/canlet

Mini Review

Common fragile sites, extremely large genes, neural development and cancer

David I Smith[a,*], Yu Zhu[a], Sarah McAvoy[a], Robert Kuhn[b]

[a]*Co-head of the Ovarian Cancer Program, Mayo Clinic Cancer Center, Mayo Clinic College of Medicine, Division of Experimental Pathology, Department of Laboratory Medicine and Pathology, Mayo Clinic, 200 First Street SW, Rochester, MN 55905, USA*
[b]*Department of Biomolecular Engineering, University of California, Santa Cruz, CA, USA*

Received 1 June 2005; accepted 6 June 2005

Abstract

Common fragile sites (CFSs) are large regions of profound genomic instability found in all individuals. They are biologically significant due to their role in a number of genomic alterations that are frequently found in many different types of cancer. The first CFS to be cloned and characterized was FRA3B, the most active CFS in the human genome. Instability within this region extends for over 4.0 Mbs and contained within the center of this CFS is the FHIT gene spanning 1.5 Mbs of genomic sequence. There are frequent deletions and other alterations within this gene in multiple tumor types and the protein encoded by this gene has been demonstrated to function as a tumor suppressor in vitro and in vivo. In spite of this, FHIT is not a traditional mutational target in cancer and many tumors have large intronic deletions without any exonic alterations. There are several other very large genes found within CFS regions including Parkin (1.37 Mbs in FRA6E), GRID2 (1.47 Mbs within 4q22.3), and WWOX (1.11 Mbs within FRA16D). These genes also appear to function as tumor suppressors but are not traditional mutational targets in cancer. Each of these genes is highly conserved and the regions spanning them are CFSs in mice. We have now examined lists of the largest human genes and found forty that span over one megabase. Many of these are derived from chromosomal bands containing CFSs. BACs within these genes are being utilized as FISH probes to determine if these are also CFS genes. Thus far we have identified the following as CFS genes: CNTNAP2 (2.3 Mbs in FRA7I), DMD (2.09 Mbs in FRAXC), LRP1B (1.9 Mbs in FRA2F), CTNNA3 (1.78 Mbs in FRA10D), DAB1 (1.55 Mbs in FRA1B), and IL1RAPL1 (1.36 Mbs in FRAXC). Although, these genes are also not traditional mutational targets in cancer they do exhibit loss of expression in multiple tumor types suggesting that they may also function as tumor suppressors. Many of the large CFS genes are involved in neurological development. Parkin is mutated in autosomal recessive juvenile Parkinsonism and deletions in mice are associated with the mouse mutant Quaking (viable). Spontaneous mouse mutants in GRID2 and DAB1 are associated with Lurcher and Reelin, respectively. In humans, alterations in IL1RAPL1 cause X-linked mental retardation and loss of WWOX is associated with Tau phosphorylation. We propose that the instability-induced alterations in these genes contribute to cancer development in a two-step process. Initial alterations will primarily occur within intronic regions, as these genes are greater than 99% intronic. These are not benign. Instead, they alter the repertoire of transcripts produced from these genes. As cancer progresses deletions will begin to encompass exons resulting in gene inactivation. These two types of alterations occurring in multiple large CFS genes may contribute significantly to

* Corresponding author. Tel.: +1 507 266 0309; fax: +1 507 266 5193.
E-mail address: smith.david@mayo.edu (D.I. Smith).

0304-3835/$ - see front matter
doi:10.1016/j.canlet.2005.06.049

the heterogeneity observed in cancer. There are also important potential linkages between normal neurological development and the development of cancer mediated by alterations in these genes.

Keywords: Common fragile sites; Cancer; Genes

1. FRA3B and FHIT

The common fragile sites (CFSs) are distinct from the rare fragile sites (RFSs) because they are found in all individuals and not in some small proportion of people who have altered DNA sequences [1]. The RFSs are expressed only after sufficient expansion of unstable repeat sequences [2]. In contrast, the CFSs are presumably unstable because of something inherent within their DNA sequence. Over 90 CFSs have been described throughout the human genome and these vary in their frequency of expression as measured by a cytogenetic assay looking for chromosomal regions with breakage/decondensation after suitable induction for CFS 'expression' [3]. The most unstable CFS region is FRA3B within chromosomal band 3p14.2 [4]. This chromosomal region is a hot-spot for deletions and other alterations in a variety of different cancers. In addition, a family was described by Cohen et al. that had a balanced reciprocal translocation within 3p14.2[t(3;8)(p14.2;q24.13)] and individuals who inherited this translocation were predisposed to develop renal cell carcinoma [5]. This suggested that an important cancer-related gene resided within this region and presumably within the FRA3B CFS.

The cloning and characterization of FRA3B revealed that it was a large region of chromosomal instability [6]. While it was initially suggested that a relatively small 300 kb region encompassed this CFS, it was later demonstrated that instability in FRA3B extended over 4.0 Mb within 3p14.2 [7]. In contrast to the rare fragile sites, there are no obvious unstable sequences which are responsible for this instability, although an analysis with the FlexStab sequence program developed in the laboratory of Dr Batsheva Kerem reveals considerable peaks indicative of sequences that could assume non-B configurations.

The identification of homozygous deletions within the FRA3B region in a variety of different cancer types led to the identification of the fragile histidine triad gene (FHIT) which had a very unusual genomic structure [8]. FHIT contains 10 small exons which together make up a 1.1 kb final processed transcript. However, these small exons span a total of 1.5 Mbs of genomic sequence. Thus, the FHIT gene covers a huge genomic stretch and greater than 99.9% of the gene is intronic sequences. Deletions and other alterations are observed in the FHIT gene in a variety of different cancers and many cancers produce aberrant FHIT transcripts of unknown significance [9].

Many observations made concerning FHIT led many to wonder if this gene was truly a tumor suppressor which was targeted during cancer development, or if it was a particularly large gene which resided within a highly unstable region. First, FHIT is not a traditional mutational target in cancer. Only a single gastric tumor was identified with a point mutation in one of the small FHIT exons [10]. While there are many tumors and tumor-derived cell lines with large deletions and homozygously deleted regions, most of the deletions occur within intronic sequences and a number of tumor cell lines were described that only had intronic alterations and produced full length wild type FHIT transcripts [11].

FHIT may not be a traditional tumor suppressor gene, but it is clear that there is an absence of expression of the Fhit protein in many different tumors and numerous premalignant lesions [9]. In addition, the FHIT $+/-$ and $-/-$ mice are more tumor prone than wild type mice after NMBA induction and these tumors can be suppressed by the addition of exogenous FHIT demonstrating that this gene does indeed function as a tumor suppressor, even if it is not mutated like a traditional tumor suppressor gene [12].

The biological function of FHIT is slowly being elucidated. Fhit hydrolyzes diadenosine tetraphosphates which are produced in cells in response to stress [13]. Over-production of Fhit results in many cells undergoing apoptosis and Fhit has been shown to

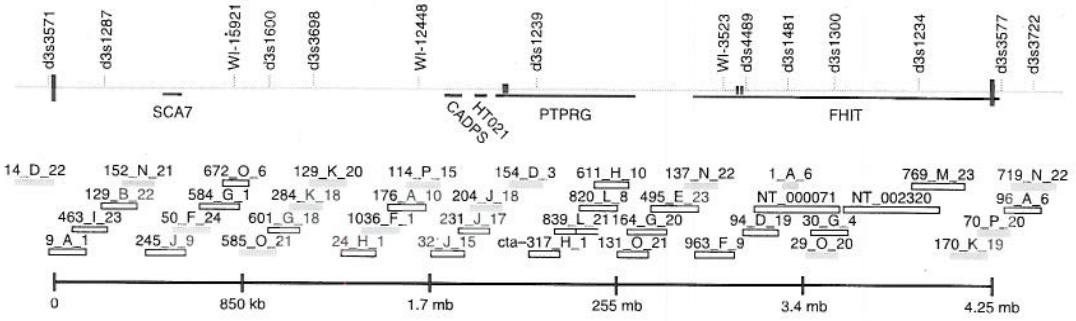


Fig. 1. Map of the 4.25 Mb FRA3B region. Included on this Figure are the genes that are localized within this region, the BAC and YAC clones which define the region, and some of the molecular markers and their location within FRA3B. The most unstable portion of this region of instability maps within the middle of the FHIT gene. Finally the FHIT gene is the most telomeric gene in this region and SCA7 is the most centromeric.

interact with a number of key proteins involved in cancer including cyclin D1 and Src [14].

Following the identification of the mouse Fhit gene, it was observed that there is considerable homology between DNA sequences surrounding the human and mouse FHIT genes, even within intronic sequences. In addition, the chromosomal region surrounding the mouse Fhit gene is also a CFS in the mouse [15]. This suggests that the very large gene and the highly unstable chromosomal region are co-conserved, possibly because they serve some function together within the cell. Fig. 1 shows the entire 4.25 Mb FRA3B region of instability and the genes, in addition to FHIT, that are localized within FRA3B.

2. FRA16D and WWOX

The second most active CFS is FRA16D (16q23.2). This chromosomal region is frequently deleted in a variety of different cancers and approximately 25% of multiple myelomas have a translocation between sequences in this region and those on chromosome 14. We localized the FRA16D CFS using a FISH-based approach with large insert YAC and BAC clones from the chromosome 16q23 region as probes. We then completely characterized the FRA16D CFS to identify the ends of this CFS region as well as the 'center' or most unstable region within FRA16D. We also physically (not electronically) identified a contig of overlapping BAC clones extending for 2.0 Mb that completely covered this CFS region [16].

The FRA16D region shares many similarities with the FRA3B region. There are no obvious sequence motifs or unstable repeats associated with either CFS region. Instability in both regions extends for at least several megabases although there is a region where the majority of the decondensation/breakage events occur, which is termed the 'center' of the CFS. In addition, both regions are associated with genes that cover very large genomic regions. In the case of FRA16D, the gene is the 1.0 Mb WWOX gene. WWOX was identified by two different groups. Rob Richards group called this gene FOR1 (for fragile oxidoreductase gene) [17], whereas the group of Manuel Aldez called the gene WWOX as this gene also has two WW domains [18]. WWOX is a 1.0 Mb gene composed of small exons that together make a 2.1 kb final processed transcript. This transcript encodes an oxidoreductase with two WW domains, hence the name WWOX is quite appropriate. This gene spans the most active region within FRA16D. There are deletions and other alterations within this large gene in a variety of different cancers, but like FHIT it is not a traditional mutational target in cancer [19]. Many cancers produce aberrant WWOX transcripts of unknown biological significance. The reintroduction of WWOX into cancer-derived cell lines that do not produce WWOX results in the inhibition of cell growth [20].

Our laboratory isolated the mouse Wox1 gene and showed that the genomic organization of the mouse Wox1 is highly conserved when compared to the human WWOX gene [21]. As was observed for FHIT,

the chromosomal region surrounding Wox1 is a CFS in the mouse [21]. Thus, FHIT and WWOX share many similarities in addition to their frequent inactivation in multiple tumor types.

The precise role that WWOX plays in the cell is also just beginning to be elucidated. WWOX has been shown to functionally associate with p73 [22], AP-2 gamma [23], and the proline rich ligand PPXY [24]. In response to stress WWOX appears to be specifically phosphorylated and in this form can bind to p53 and induce apoptosis [25]. Thus, for FHIT and WWOX there is a potential association with cellular responses to stress.

The down-regulation of WWOX induces Tau phosphorylation in vitro [26]. It was therefore suggested that WWOX potentially plays a role in Alzheimer's disease. This is the first indication that any of the large CFS genes would be involved in neurological development or neurodegeneration.

3. FRA6E and Parkin

A number of different methods have been used to localize and characterize CFS regions. Both FRA3B and FRA16D were localized by using large insert clones as FISH probes to triangulate and eventually uncover each CFS region. This strategy was also utilized to localize a number of other CFS regions including FRA7G [27], FRAXB [28], and FRA2G [29]. An alternative strategy to localize many other CFS regions was based upon the observation that human papillomaviruses, HPV16 and HPV18, were preferentially integrated into CFS regions in cervical tumors [30,31]. Over half of the sites of viral integration identified turned out to be CFS regions. By identifying the DNA sequences immediately adjacent to the sites of HPV integration in different cervical tumors we were able to localize 21 CFS regions. A third successful strategy that identified six previously uncharacterized CFS regions was based upon the identification of genes whose expression was consistently down-regulated during the development of ovarian cancer and testing large insert clones spanning those genes to determine if the genes were derived from within a CFS region [32]. This strategy identified tsg101, ARH1, TPM1, and IGF2R, PLG and SLCC22A3 as CFS genes. IGF2R, PLG and SLC22A3 were all localized on the proximal end of the FRA6E (6q26) CFS. We then completely defined the FRA6E region with BAC clones and found that instability within this region extended for 3.6 Mbs [33]. Spanning the distal half of this CFS, and the active 'center' of the FRA6E CFS is a third extremely large CFS gene, Parkin [33].

Parkin was first identified as a mutational target in some patients with autosomal recessive juvenile Parkinsonism (ARJP) [34]. As a result of this most of the work characterizing Parkin has focused on the role of Parkin in neural cells and its role in the neurodegeneration that occurs in ARJP patients. Parkin has been shown to interact with a number of key neural proteins including synuclein [35]. Less is known about Parkin and its role in epithelial cells, although N-myc has been shown to regulate Parkin expression [36].

Parkin spans 1.36 Mbs and is comprised of 11 small exons that together comprise a final processed transcript of 2.3 kbs [33]. Parkin expression was down-regulated in 60% of primary ovarian tumors analyzed and many tumors had alternative Parkin transcripts of unknown significance [33,37]. The biological significance of loss of Parkin expression is unknown but we have demonstrated that the re-introduction of Parkin into cell lines that did not express it resulted in those cells being more sensitive to apoptotic induction [33].

4. GRID2 and FRA4?? (4q22)

GRID2 is another extremely large gene (1.46 Mbs) and Michelle Debatisse and co-workers demonstrated that this gene is localized within a CFS region in both humans and mice [38]. In addition, the mouse gene is a hot-spot for spontaneous deletions resulting in the mouse neurological mutant Lurcher. There is also considerable homology between the mouse and human GRID2 genes even within the introns. Recurrent deletions of subregions of band 4q22 has been described in human hepatocellular carcinomas suggesting that GRID2 may also play a role in hepatic carcinogenesis [39]. The full size of the region of instability surrounding the FRA4?? CFS is over 7 Mb and in addition to GRID2 there is another extremely large gene KIAA1680 (1.47 Mbs), located within the region [38].

5. Not all CFS regions are associated with extremely large genes

Our observations with several of the most active of the CFS regions and their association with large genes are not applicable to all CFS regions. Indeed more of the characterized CFS regions are not associated with extremely large genes. This includes FRA7G [27], FRAXB [28], FRA7H [40], FRA2G [29] and FRA6F [41]. Some of the CFS regions contain several smaller genes, such as FRA7G, FRAXB, FRA2G and FRA6F, while the 300 kb region spanned by FRA7H is not associated with any genes [40].

6. The largest human genes

In spite of the fact that not all CFS regions are associated with genes that span vast genomic stretches, we were interested in whether there were other very large genes that could also be derived from CFS regions. We were also curious how large FHIT, Parkin and WWOX were relative to the largest human genes. We obtained lists of the largest known human genes from Dr Robert Kuhn (UCSC Database) and after carefully curating those lists to remove redundant genes discovered that there were 40 human genes that spanned greater than 1.0 Mb and another 200 that spanned between 500 kb and 1.0 Mb of genomic sequence. Our three large CFS genes were actually the 10th (FHIT), 17th (Parkin) and 33rd (WWOX) largest human genes. Table 1 shows each of the 40 human genes that span greater than 1.0 Mb of genomic sequence. Also included on this Table is the chromosomal location of each gene, the number of exons for each gene and the size of the final processed transcript.

7. Many of the largest human genes are localized to chromosomal regions that contain CFSs

An examination of the Table 1 reveals that many of the largest known human genes do indeed map to chromosomal bands that contain a CFS. Although only a few of the CFS regions have been completely characterized we do know the approximate location of 20 additional CFS regions. These were delineated either by the identification of a viral integration site in a cervical tumor or because gene(s) within those regions frequently lost expression in primary ovarian tumors. Although we do not know where these 20 CFS regions begin, center and end, we do have information about a single BAC from within each region and the frequency that the BAC hybridized proximal, distal or crossing its respective CFS.

We examined the list of 240 genes that span over 500 kb of genomic sequence and then scanned the regions surrounding each of the 31 localized CFSs to determine if any of these genes were within or close to the CFS regions. We found that CNTNAP2, the largest known human gene (2.3 Mb) is actually localized within the FRA7I CFS. We also discovered that LRP1B, a 1.9 Mb gene, is located within a CFS region, FRA2F (2q22.1). A third large CFS gene was identified because it was frequently inactivated in ovarian cancers, PDGFRA (in FRA4B) [32].

Many of the large CFS genes were found to reside immediately adjacent to other very large genes. In FRA3B (see Fig. 1) the 720 kb PTPRG gene is immediately centromeric to FHIT. In FRA6E, Parkin is centromeric to the 800 kb PACRG gene. GRID2 in FRA4?? is just telomeric of KIAA1680. We then examined the list of 240 genes that spanned greater than 500 kb to determine if other large genes were similarly clustered. This revealed that there was a decidedly non-random distribution for the largest human genes with many of them residing within chromosomal regions that contain multiple large genes. In addition, it suggested several likely CFS candidates. For example, DMD is localized immediately adjacent to IL1RAPL1 in chromosomal band Xp21.1.

8. Examining very large genes as possible CFS genes

We used several criteria to determine which large genes to test as potential CFS genes. Since we were particularly interested in genes that could play an important role in the development of cancer, we chose the 1.2 Mb deleted in colorectal cancer (DCC) gene, as well as the 677 kb RAD51L1 gene, which is the human homolog of the bacterial recA gene. The DCC gene is derived from 18q21.1 and the Rad51L1 gene is

Table 1
Very large human genes and their chromosomal localization

Largest human genes					
	Gene name	Chromosome	Size	Exons/FPT	Closest CFS
1	CNTNAP2	7q35	2304258	25/8107	FRA7I
2	DMD	Xp21.1	2092287	79/13957	FRAXC
3	CSMD1	8p23.2	2056709	70/11580	FRA8B
4	LRP1B	2q22.1	1900275	91/16556	FRA2F
5	CTNNA3	10q21.3	1775996	18/3024	FRA10D
6	NRXN3	14q24.3	1691449	21/6356	FRA14C
7	A2BP	16p13.2	1691217	16/2279	
8	DAB-1	1p32.3	1548827	21/2683	FRA1B
9	PDE4D	5q11.2	1513407	17/2465	
10	FHIT	3p14.2	1499181	9/1095	FRA3B (3p14.2)
11	KIAA1680	4q22.1	1474315	11/5833	FRA4??
12	GPC5	13q31.3	1468199	8/2588	FRA13D?
13	GRID2	4q22.3	1467842	16/3024	FRA4??
14	DLG2	11q14.1	1463760	23/3071	FRA11F
15	AIP1	7q21.11	1436474	21/6795	FRA7E
16	DPP10	2q14.1	1402038	26/4905	
17	Parkin	6q26	1379130	12/2960	FRA6E (6q26)
18	ILIRAPL1	Xp21.2	1368379	11/2722	FRAXC
19	PRKG1	10q21.1	1302704	18/2213	FRA10C
20	EB-1	12q23.1	1248678	26/3750	FRA12C
21	CSMD3	8q23.2	1213952	69/12486	FRA8C
22	IL1RAPL2	Xq22.3	1200827	11/2985	
23	AUTS2	7q11.22	1193536	19/5972	FRA7J
24	DCC	18q21.1	1190131	29/4608	FRA18B
25	GPC6	13q31.3	1176822	9/2731	FRA13D
26	CDH13	16q23.2	1169565	15/3926	FRA16D distal
27	ERBB4	2q34	1156473	28/5484	FRA2I
28	ACCN1	17q11.2	1143718	10/2748	
29	CTNNA2	2p12	1135782	18/3853	FRA2E
30	WD repeat	2q24	1126043	16/2132	
31	DKFZp686H	11q25	1117478	8/6830	FRA11G
32	PTPRT	20q12	1117144	32/12680	
33	WWOX	16q23.2	1113013	9/2264	FRA16D (16q23.2)
34	NRXN1	2p16.3	1109951	21/8114	FRA2D
35	IGSF4D	3p12.1	1109105	10/3315	
36	CDH12	5p14.3	1102578	15/4167	FRA5E
37	PAR3L	2q33.3	1069815	23/4176	FRA2I
38	PTPRN2	7q36.3	1048712	22/4735	FRA7I
39	SOX5	12p12.1	1030095	18/4492	
40	TCBA1	6q22.31	1021499	8/3183	FRA6F

derived from 14q24.1. We treated lymphocytes with 0.4 uM aphidicolin for 24 h and prepared metaphase chromosome preparations. We then obtained BAC clones that spanned the 5′ ends of DCC and Rad51L1, fluorescently labeled them, and used them as FISH probes against the aphidicolin-treated metaphase preparations. We then analyzed at least 20 metaphases with good discernible breakage at the FRA14C (14q24.2) and FRA18B (18q21.3) CFSs, respectively. We found that the RAD51 BAC always hybridized distal to breakage within the 14q24.1 FRA14C CFS, and that the DCC BAC always hybridized proximal to breakage within FRA18B. Hence, these two genes are not derived from within CFS regions.

We next examined IL1RAPL1 from Xp21.3 as this 1.3 Mb gene was localized immediately proximal of the 2.09 Mb DMD gene. Unfortunately, the FRAXC CFS is expressed at very low frequencies, hence we had to examine many metaphases until we had found even a few with breakage in the FRAXC region. In preliminary studies we have demonstrated that IL1RAPL1 is located within this CFS region and we are currently testing whether DMD is also within the unstable region.

Four other large genes that were analyzed were CTNNA3 (alpha-T catenin), DAB1 (the human homolog of Drosophila disabled), RORA (the orphan retinoic acid receptor alpha), and LARGE. BAC clones spanning the 5′ portions of each of these genes were used as FISH based probes and we found that each of these genes were also CFS genes, localizing to FRA10D (CTNNA3), FRA1B (DAB1), FRA15A (RORA) and FRA22B (LARGE).

We have now demonstrated that 6 of the 10 largest human genes are derived from within CFS regions. Two of the genes, A2BP and PDE4D are not derived from within CFS regions as there are no CFSs on the short arm of chromosome 16 or anywhere near 5q11.2. We have not yet tested the 8p23.2 CSMD1 gene or the 14q24.3 NRXN3 gene, to determine if they localize within the FRA8B and FRA14C CFSs, respectively.

What proportion of the CFSs are associated with very large genes? The complete definition of a CFS region is an extensive effort that requires multiple BAC clones covering the full region of instability and this analysis has only been done for eight of the CFS regions, FRA3B, FRA16D, FRA6E, FRA6F, FRA2G, FRA9E, FRA4??, and FRAXB. Four of these CFS regions are associated with an extremely large gene. However, an additional 23 CFS regions have at least been localized. We therefore examined a 5–10 Mb region around each localized CFS and then searched these regions for any extremely large genes. We found that only 8 of the 23 localized CFS regions were associated with large genes. Taken together, we have found that 12 of the 31 CFS regions are associated with large genes. We can thus roughly estimate that as many as 30 of the 90 known CFS regions will contain extremely large genes like FHIT and Parkin.

The large genes that have now been definitely found to reside within a CFS region include DAB1 (FRA1B, 1p32.3), LRP1B and ARHGAP15 (FRA2F, 2q22.1), FHIT and PTPRG (FRA3B, 3p14.2), GRID2 and KIAA1680 (in FRA4??, 4q22.1), Parkin and PACRG (FRA6E, 6q26), CNTNAP2 (FRA7I, 7q35), CTNNA3 (FRA10D, 10q21.3), DLG2 (FRA11F, 11q14.1), RORA (FRA15A, 15q22.2), WWOX (FRA16D, 16q23.2), LARGE (FRA22B, 22q12.3) and IL1RAPL1 and DMD (FRAXC, Xp21.2).

9. Large genes and neurological development

The first large CFS gene that was identified to be involved in neurological development was Parkin, which is mutated in some patients with autosomal recessive juvenile Parkinsonism. In mice there is a spontaneous deletion of Parkin and the immediately distal PARCG gene that results in the Quaker (viable) phenotype [42].

A second large CFS gene is the delta2 glutamate receptor gene (GRID2) which is localized within a chromosomal region in the mouse where there are frequent spontaneous rearrangements. Deletion of GRID2 results in the mouse phenotype Lurcher, which is associated with ataxia as a result of selective, cell-autonomous and apoptotic death of cerebellar Purkinje cells [43].

RORA encodes an orphan retinoic receptor which is involved in the control of circadian rhythm [44]. In addition, RORA binds to the hypoxia inducible factor and thus may be involved in cellular responses to hypoxia and other stresses [45]. Deletion of RORA in the mouse results in another mouse neurological mutant, Staggerer, which is associated with tremors, body imbalance, small size and they generally die between 3 and 4 weeks [46]. In addition to their small size, the Staggerer mice have fewer ectopically localized Purkinje cells.

A number of the other large CFS genes also appear to play important roles in neurological development. The largest human gene, CNTNAP2 (2.3 Mbs) is disrupted in a family with Gilles de la Tourette syndrome [47]. The low density lipoprotein receptor-related protein 1B (LRP1B) retains beta-amyloid at the cell surface and reduces amyloid-beta production [48]. DMD is mutated in patients with Duchene Muscular Dystrophy and the tightly linked IL1RAPL1 gene is associated with X-linked mental retardation

[49]. Many of the other large CFS genes have also been found to be important in neurological development.

10. Role of the CFSs and the large genes contained within them in normal cells

Since the genes within the CFS regions are highly susceptible to genomic instability especially within developing cancer cells, most of the studies on the CFS genes have been in the context of what role they could play in cancer development. A number of these studies have revealed that many of the large CFS genes do play important roles in cancer development and several of the large CFS genes appear to function as tumor suppressors, even if they are not traditional mutational targets in cancer.

However, little work has been done to determine the function of the CFSs and the large genes contained within them in the normal cell. The observation that these extremely large genes residing within some of the most unstable chromosomal regions are highly evolutionarily conserved even within intronic regions and that the CFS and the large genes are co-conserved suggests that these genes and the unstable region share some function within normal cells.

One observation that is interesting is that so many of the large CFS genes appear to function as stress responders which leads us to the hypothesis that the unstable CFSs and their co-conserved large genes function together as a stress response system within cells. However, it still remains very unclear how this system functions or works. One possibility is that somehow the CFS regions are able to transduce cellular stresses into different transcripts being produced from the large CFS genes. An alternative is that the very large introns produce transcripts which somehow regulate the expression of the large CFS genes. In order to determine if either of these are likely we are initiating experiments to characterize the transcript isoforms that are made from these large genes in response to cellular stress. We are also exploring ways of examining the large introns for RNA transcripts that are produced and regulated in response to stress.

Two papers were recently published whose results could provide additional support for our hypotheses. The first was done by the group of Tom Gingeras and co-workers at Affymetrix. Utilizing oligonucleotide tiling arrays for human chromosomes 21 and 22 they determined that there were 10 times the number of transcripts produced from these two chromosomes than would be anticipated from the determined number of genes on those chromosomes. This suggests that there are many more RNA species encoded within the genome than previously anticipated. The second finding comes from the group of Ted Krontiris at the City of Hope. They have been searching for prostate cancer susceptibility alleles and now using high resolution SNP analysis have localized one such allele within one of the large introns of the FHIT gene. This suggests that changes in DNA sequences within the introns of one of these genes could have dramatic effects, again supporting the hypothesis that the large introns may be producing RNA species whose function is to regulate gene expression.

In conclusion we have now found that almost half of the 20 largest human genes are derived from within CFS regions. Since these genes are localized within some of the most unstable chromosomal regions in the genome, they are uniquely susceptible to increases in genomic instability such as occurs during the development of many different cancers. However, these genes are all greater than 99.8% intronic, thus many of the alterations that occur within these genes occur within the introns. However, these alterations are not benign as they may alter important transcripts whose function is to regulate the expression of the CFS gene. Finally, many of the large CFS genes appear to function both as part of a stress response system and also in normal neurological development.

We are continuing our studies to characterize the large highly unstable CFS regions and the extremely large genes contained within them. Our work is focused on characterizing how these regions and the genes could function together as a stress response system within normal cells. In addition, since we estimate that there are as many as 30 large CFS genes within the genome, we are also examining the role that alterations within these genes could play in the development of many different types of cancer.

References

[1] T.W. Glover, Instability at chromosomal fragile sites, Recent Results Cancer Res. 154 (1998) 185–199.

[2] G.R. Sutherland, E. Baker, R.I. Richards, Fragile sites still breaking, Trends Genet. 14 (1998) 501–506.

[3] D.I. Smith, H. Huang, L. Wang, Common fragile sites and cancer (review), Int. J. Oncol. 12 (1998) 187–196.

[4] D. Smeets, J. Scheres, T. Hustinx, The most common fragile site in man is 3p14, Hum. Genet. 72 (1986) 215–220.

[5] A.J. Cohen, F.P. Li, S. Berg, D.J. Marchetto, S. Tsai, S.C. Jacobs, et al., Hereditary renal-cell carcinoma associated with a chromosomal translocation, New Eng. J. Med. 301 (1979) 592–595.

[6] W. Paradee, C.M. Wilke, L. Wang, R. Shridhar, C.M. Mullins, A. Hoge, et al., A 350 Kb cosmid contig in 3p14.2 that crosses the t(3;8) hereditary renal cell carcinoma translocation breakpoint and 17 aphidicolin-induced FRA3B breakpoints, Genomics 35 (1996) 87–93.

[7] N.A. Becker, E.C. Thorland, S.R. Denison, L.A. Phillips, D.I. Smith, Evidence that instability within the FRA3B region extends four megabases, Oncogene 21 (2002) 8713–8722.

[8] T. Druck, P. Hadaczek, T.B. Fu, M. Ohta, Z. Siprashvili, R. Baffa, et al., Structure and expression of the human FHIT gene in normal and tumor cells, Cancer Res. 57 (1997) 504–512.

[9] T. Druck, L. Berk, K. Huebner, FHITness and cancer, Oncol. Res. 10 (1998) 341–345.

[10] A. Gemma, K. Hagiwara, Y. Ke, L.M. Burke, M.A. Khan, M. Nagashima, et al., FHIT mutations in primary gastric cancer, Cancer Res. 57 (1997) 504–512.

[11] M.M. LeBeau, H. Drabkin, T.W. Glover, R. Gemmill, F.V. Rassool, T.W. McKeithan, et al., An FHIT tumor suppressor gene?, Genes Chromosom. Cancer 21 (1998) 281–289.

[12] H. Ishii, A. Vecchione, L.Y. Fong, N. Zanesi, F. Trapasso, Y. Furukawa, et al., Cancer prevention and therapy in a preclinical mouse model: impact of FHIT viruses, Curr. Gene. Ther. 4 (2004) 53–63.

[13] L.L. Kisselev, J. Justesen, A.D. Wolfson, L.Y. Frolova, Diadenosine oligophosphates (Ap(n)A), a novel class of signaling molecules, Fed. Euro. Biochem. Soc. Lett. 427 (1998) 157–163.

[14] A. Matsuyama, C.M. Croce, K. Huebner, Common fragile genes, Eur. J. Histochem. 48 (2004) 29–36.

[15] A. Matsuyama, T. Shiraishi, F. Trapasso, T. Kuroki, H. Alder, M. Mori, et al., Fragile site orthologs FHIT/FRA3B and Fhit/Fra14A2: evolutionarily conserved but highly recombinogenic, Proc. Natl. Acad. Sci. USA 100 (2003) 14988–14993.

[16] K.A. Krummel, L.R. Roberts, M. Kawakami, T.W. Glover, D.I. Smith, The characterization of the common fragile site FRA16D & its involvement in multiple myeloma translocations, Genomics 69 (2000) 37–46.

[17] K. Ried, M. Finnis, L. Hobson, M. Mangelsdorf, S. Dayan, J.K. Nancarrow, et al., Common chromosomal fragile site FRA16D sequene: identification of the FOR gene spanning FRA16D and homozygous deletions and translocation breakpoints in cancer cells, Hum. Mol. Genet. 9 (2000) 1651–1663.

[18] A.K. Bednarek, K.J. Laflin, R.L. Daniel, Q. Liao, K.A. Hawkins, C.M. Aldaz, WWOX, a novel WW domain-containing protein mapping to human chromosome 16q23.3-24.1, a region frequently affected in breast cancer, Cancer Res. 60 (2000) 2140–2145.

[19] K. Driouch, H. Prydz, R. Monese, H. Johansen, R. Lidereau, E. Frengen, Alternative transcripts of the candidate tumor suppressor gene, WWOX, are expressed at high leveles in human breast tumors, Oncogene 21 (2002) 1832–1840.

[20] A.K. Bednarek, C.L. Keck-Waggoner, R.L. Daniel, K.J. Laflin, P.L. Bergsagel, K. Kiguchi, et al., WWOX, the FRA16D gene, behaves as a suppressor of tumor growth, Cancer Res. 61 (2001) 8068–8073.

[21] K.A. Krummel, S.R. Denison, E. Calhoun, L.A. Phillips, D.I. Smith, The common fragile site FRA16D and its associated gene WWOX are highly conserved in the mouse at Fra8E1, Genes Chromosom. Cancer 34 (2002) 154–167.

[22] R.I. Aqeilan, Y. Pekarsky, J.J. Herrero, A. Palamarchuk, J. Letofsky, T. Druck, et al., Functional association between Wwox tumor suppressor protein and p73, a p53 homolog, Proc. Natl. Acad. Sci. USA 101 (2004) 4401–4406.

[23] R.I. Aqeilan, A. Palamarchuk, R.J. Weigel, J.J. Herrero, Y. Pekarsky, C.M. Croce, Physical and functional interactions between the Wwox tumor suppressor protein and the AP-2 gamma transcription factor, Cancer Res. 64 (2004) 8256–8261.

[24] J.H. Ludes-Meyers, H. Kil, A.K. Bednarek, J. Drake, M.T. Bedford, C.M. Aldaz, WWOX binds the specific praline-rich ligand PPXY: identification of candidate interacting proteins, Oncogene 23 (2004) 5049–5055.

[25] N.S. Chang, J. Doherty, A. Ensign, J. Lewis, J. Heath, L. Schultz, et al., Molecular mechanisms underlying WOX1 activation during apoptotic and stress responses, Biochem. Pharmacol. 66 (2003) 1347-13.

[26] C.I. Sze, M. Su, S. Puzazhenthi, P. Jambal, L.J. Hsu, J. Heath, et al., Down-regulation of WW domain-containing oxidoreductase induces Tau phosphorylation in vitro: a potential role in Alzheimers disease, J. Biol. Chem. 279 (2004) 30498–30506.

[27] H. Huang, C. Qian, R.B. Jenkins, D.I. Smith, Fish mapping of YAC clones at human chromosomal band 7q31.2: identification of YACs spanning FRA7G within the common region of LOH in breast and prostate cancer, Genes Chromosom. Cancer 21 (1998) 152–159.

[28] M.F. Arlt, D.E. Miller, D.G. Beer, T.W. Glover, Molecular characterization of FRAXB and comparative common fragile site instability in cancer cells, Genes Chromosom. Cancer 33 (2002) 82–92.

[29] M.Z. Limongi, F. Peilliccia, A. Rocchi, Characterization of the human common fragile site FRA2G, Genomics 81 (2003) 93–97.

[30] E.C. Thorland, S.L. Myers, B.S. Gostout, D.I. Smith, Common fragile sites are preferential targets for HPV16 integrations in cervical tumors, Oncogene 22 (2003) 1225–1237.

[31] M.J. Ferber, E.C. Thorland, A.A. Brink, A.K. Rapp, L.A. Phillips, R. McGovern, et al., Preferential integration of human papillomavirus type 18 near the c-myc locus in cervical carcinoma, Oncogene 22 (2003) 7233–7242.

[32] S.R. Denison, N.A. Becker, M.J. Ferber, L.A. Phillips, K.R. Kalli, J. Lee, et al., Transcriptional profiling reveals that several common fragile-site genes are downregulated in ovarian cancer, Genes Chromosom. Cancer 34 (2002) 406–415.

[33] S.R. Denison, F. Wang, N.A. Becker, B. Schule, N. Kock, L.A. Phillips, et al., Alterations in the common fragile site gene Parkin in ovarian and other cancers, Oncogene 22 (2003) 8370–8378.

[34] T. Kitada, S. Asakawa, N. Hattori, H. Matsumini, Y. Yamarua, S. Minoshima, et al., Mutations in the parkin gene cause autosomal recessive juvenile parkinsonism, Nature 392 (1998) 605–608.

[35] K.K. Chung, Y. Zhang, K.L. Lim, Y. Tanaka, H. Huang, J. Gao, et al., Parkin ubiquitinates the alpha-synuclein-interacting protein, synphilin-1: implications for Lewy body formation in Parkinson disease, Nat. Med. 7 (2001) 1144–1150.

[36] A.B. West, G. Kapatos, C. O'Farrell, F. Gonzalez-de-Chavez, K. Chiu, M.J. Farrer, et al., J. Biol. Chem. 279 (2004) 28896–28902.

[37] F. Wang, S. Denison, J.P. Lai, L.A. Phillips, D. Montoya, N. Kock, et al., Parkin gene alterations in hepatocellular carcinoma, Genes Chromosome. Cancer 40 (2004) 85–96.

[38] L. Rozier, E. El-Achkar, F. Apiou, M. Debatisse, Characterization of a conserved aphidicolin-sensitive common fragile site at human 4q22 and mouse 6C1: possible association with an inherited disease and cancer, Oncogene 23 (2004) 6872–6880.

[39] O. Bluteau, J.C. Beaudoin, P. Pasturaud, J. Belghiti, D. Franco, P. Biolac-Sage, et al., Specific association between alcohol intake, high grade of differentiation and 4q34-q35 deletions in hepatocellular carcinomas identified by high resolution allelotyping, Oncogene 21 (2002) 1225–1232.

[40] D. Mishmar, A. Rahat, S.W. Scherer, G. Nyakatura, B. Hinzmann, Y. Kohwi, et al., Molecular characterization of a common fragile sie (FRA7H) on human chromosome 7 by the cloning of a simian virus 40 integration site, Proc. Natl. Acad. Sci. USA 95 (1998) 8141–8146.

[41] C. Morelli, E. Karayianni, C. Magnanini, A.J. Mungall, E. Thorland, M. Negrini, et al., Cloning and characterizatjon of the common fragile site FRA6F harboring a replicative senescence gene and frequently deleted in human tumors, Oncogene 21 (2002) 7266–7276.

[42] D. Lorenzetti, B. Antalffy, H. Bogel, J. Novesroske, D. Armstrong, M. Justice, The neurological mutant quaking (viable) is Parkin deficient, Mamm. Genome 15 (2004) 210–217.

[43] M.L. Doughty, P.L. De Jagel, S.J. Korsmeyer, N. Heintz, Neurodegeneration in Lurcher mice occurs via multiple cell death pathways, J. Neurosci. 20 (2000) 5339–5345.

[44] T.K. Sato, S. Panda, L.J. Miraglia, T.M. Reyes, R.D. Rudic, P. McNamara, et al., A functional genomics strategy reveals Rora as component of the mammalian circadian clock, Neuron 43 (2004) 527–537.

[45] C. Chauvet, B. Bois-Joyeux, E. Berra, J. Pouyssegur, J.L. Danan, The gene encoding human retinoic acid-receptor-related orphan receptor alpha is a target for hypoxia-inducible factor 1, Biochem. J. (2004) 79–85.

[46] I. Dussault, D. Fawcett, A. Matthyssen, J.A. Bader, V. Giguere, Orphan nuclear receptor ROR alpha-deficient mice display the cerebellar defects of staggerer, Mech. Dev. 70 (1998) 147–153.

[47] A.J. Verkerk, C.A. Mathews, M. Joosse, B.H. Eussen, P. Heutink, B.A. Oostra, Tourette syndrome association international consortium for genetics. CNTNAP2 is disrupted in a family with Giles de la Tourette syndrome and obsessive compulsive disorder, Genomics 82 (2003) 1–9.

[48] J.A. Cam, C.V. Zerbinatti, J.M. Knisely, S. Hecimovic, Y. Li, G. Bu, The low density lipoprotein receptor-related protein 1B retains beta-amyloid precursor protein at the cell surface and reduces amyloid-beta peptide production, J. Biol. Chem. 279 (2004) 29639–29646.

[49] Y.H. Zhang, B.L. Huang, K.K. Niakan, L.L. McCabe, E.R. McCabe, K.M. Dipple, IL1RAPL1 is associated with mental retardation in patients with complex glycerol kinase deficiency who have deletions extending telomeric of DAX1, Hum. Mutat. 24 (2004) 273.

Oncogene (2006) 00, 1–8

www.nature.com/onc

ORIGINAL ARTICLE

RORA, a large common fragile site gene, is involved in cellular stress response

Y Zhu[1,3], S McAvoy[1,3], R Kuhn[2] and DI Smith[1]

[1]*Department of Laboratory Medicine and Pathology, Division of Experimental Pathology, Mayo Clinic College of Medicine, Rochester, MI, USA and* [2]*Department of Biomolecular Science and Engineering, University of California, Santa Cruz, CA, USA*

Common fragile sites (CFSs) are large genomic regions present in all individuals who are highly unstable and prone to breakage and rearrangement, especially in cancer cells with genomic instability. Eight of the 90 known CFSs have been precisely defined and five of these span genes that extend from 700 kb to over 1.5 Mb of genomic sequence. Although these genes reside within some of the most unstable chromosomal regions in the human genome, they are highly conserved evolutionarily. These genes are targets for large chromosomal deletions and rearrangements in cancer and are frequently inactivated in multiple tumor types. There is also an association between these genes and cellular responses to stress. Based upon the association between large genes and CFSs, we began to systematically test other large genes derived from chromosomal regions that were known to contain a CFS. In this study, we demonstrate that the 730 kb retinoic acid receptor-related orphan receptor alpha (RORA) gene is derived from the middle of the FRA15A (15q22.2) CFS. Although this gene is expressed in normal breast, prostate and ovarian epithelium, it is frequently inactivated in cancers that arise from these organs. RORA was previously shown to be involved in the cellular response to hypoxia and here we demonstrate changes in the amount of RORA message produced in cells exposed to a variety of different cellular stresses. Our results demonstrate that RORA is another very large CFS gene that is inactivated in multiple tumors. In addition, RORA appears to play a critical role in responses to cellular stress, lending further support to the idea that the large CFS genes function as part of a highly conserved stress response network that is uniquely susceptible to genomic instability in cancer cells.
Oncogene (2006) **0,** 000–000. doi:10.1038/sj.onc.1209314

Q1 **Keywords:** ▮; ▮; ▮;

Correspondence: Professor DI Smith, Department of Laboratory Medicine and Pathology, Division of Experimental Pathology, Mayo Clinic College of Medicine, 200 First Street, S.W., Rochester, MN 55905, USA.
E-mail: smith.david@mayo.edu
[3]These authors contributed equally to this work.
Received 20 July 2005; revised 3 November 2005; accepted 3 November 2005

Introduction

Fragile sites, which consist of rare fragile sites (RFSs) and common fragile sites (CFSs), are specific chromosomal loci that non-randomly exhibit gaps or breaks in response to specific culture conditions or exposure to certain chemical agents. In contrast to the RFSs, which are found in less than 5% of the population and whose instability is associated with expansion of some repeat sequence, the CFSs are present in all individuals and are not found to be associated with any simple unstable repeat sequences (Sutherland and Richards, 1995).

So far, a total of 90 CFS regions have been identified throughout the human genome (Buttel *et al.*, 2004). CFSs are highly unstable and recombinogenic regions of the genome. They are preferential sites of sister chromatid exchange, translocations, deletions, intrachromosomal gene amplification and integration of DNA from tumor-associated viruses. As the CFS regions are so unstable, it has been presumed that genes residing within these regions are frequently altered by the deletions and rearrangements that occur in genetically unstable cancer cells. Therefore, it has been proposed that CFSs, and the genes located within them, play a mechanistic role in the initiation or progression of human cancers (Arlt *et al.*, 2003; Buttel *et al.*, 2004).

At this point, only a few of the CFSs have been fully characterized. The most unstable regions include FRA3B (3p14.2) and FRA16D (16q23.2) (Smith *et al.*, 1998; Sutherland *et al.*, 1998). These CFSs are large regions of genomic instability spanning multiple megabases. Contained within the most unstable regions within these CFSs are genes that themselves span very large genomic regions. For example, the total size of the most unstable CFS (FRA3B) is over 4 Mb and spanning the most unstable portion of this CFS is the extremely large FHIT gene (Huebner *et al.*, 1997; Becker *et al.*, 2002). The genomic size of FHIT is over 1.5 Mb, but it is comprised of only nine small exons that make a 1095 bp final transcript. Surprisingly, FHIT is not a traditional mutational target in human cancer, as only a single gastric tumor has been identified with a point mutation in this gene (Gemma *et al.*, 1997). However, there are frequent deletions and other alterations within this large gene in many different types of cancer (Druck *et al.*, 1997). In addition, many primary tumors do not express

Gml : Ver 6.0
Template: Ver 6.1

Journal: ONC ☐ Disk used Despatch Date: 14/12/2005
Article : npg_onc_2005-00995 Pages: 1–8 OP: Deeparani ED: Sudhir

the Fhit protein (Becker-Andre *et al*., 1993; Becker *et al*., 2002).

WWOX, which spans 1.0 Mb of genomic sequence in 16q23.2, has been identified within FRA16D, another chromosomal region frequently deleted in multiple cancers (Ludes-Meyers *et al*., 2003). WWOX, like FHIT, is not a traditional mutational target in cancer, but there are deletions and other alterations within the highly unstable region surrounding WWOX, and it is also frequently inactivated in many different tumors (Finnis *et al*., 2005). In response to specific types of stress, WWOX is specifically phosphorylated, and phosphorylated WWOX then binds to p53, translocates to the nucleus and induces apoptosis (Chang *et al*., 2003). Functional studies have demonstrated that both FHIT and WWOX may work as tumor suppressors (Bednarek *et al*., 2001; Dumon *et al*., 2001).

These large CFS genes are remarkably evolutionarily conserved not only in the small exons but also in the large intronic regions. In addition, the chromosomal regions spanning these large genes in the mouse are also CFSs (Krummel *et al*., 2002; Matsuyama *et al*., 2003); thus, the large genes and highly unstable regions are co-conserved, suggesting that together they serve some important cellular function. As the CFS regions and the large size CFS genes were associated with human cancer development, we began to systematically test a number of large genes that were derived from chromosomal regions known to contain a CFS. It was found that the retinoid-related orphan receptor alpha (RORA) gene, which encodes an ROR and is 730 kb size, is located in the middle of the FRA15A (15q22.2) CFS.

RORA encodes an ROR, which functions as an evolutionarily related transcription factor and belongs to the steroid hormone receptor superfamily (Jetten, 2004). It consists of three members: RORA (also named NRF1 by the Nuclear Receptor Nomenclature Committee), RORB (NRF2) (Carlberg *et al*., 1994) and RORC (NRF3, or TOR) (Hirose *et al*., 1994). It has been demonstrated that these receptors are critical in the regulation of a number of physiological processes. However, the involvement of individual RORs in possible physiological function *in vivo* is still poorly understood. The RORA gene produces four isoforms (RORA1–RORA4), which only differ in their N-terminal regions and demonstrate distinct DNA-binding and transactivation properties (Becker-Andre *et al*., 1993; Giguere *et al*., 1994; Matysiak-Scholze and Nehls, 1997). In the RORA-mutated mouse, *staggerer*, there are specific cerebellar abnormalities, showing that this nuclear receptor plays a critical role in the development of the cerebellum (Hamilton *et al*., 1996). RORA has also been suggested to be involved in lipid metabolism, to possess immunomodulatory activity and to mediate the antiarthritic properties of a class of thiazolidinediones (Missbach *et al*., 1996). ROREs (ROR response elements), to which RORA protein binds, have been identified in the promoter region of cell cycle-related genes, such as those of the cyclin-dependent kinase (CDK) inhibitor $p21^{WAF1/CIP1}$ (Schrader *et al*., 1996), and of cyclin A, as well as in the promoter of N-myc (Lee *et al*., 1984; Nau *et al*., 1986), a gene whose amplification appears to be related to the development of several tumors. It has been reported that ligand-induced activation of RORA significantly reduces the growth of the murine colon 38 adenocarcinoma (Pawlikowski *et al*., 1999). Taken together, these observations suggest that unlike the other two members in this family, RORA might be involved in the regulation of cell growth and tumorigenesis.

The goal of this study was to determine if RORA was another large CFS gene and if it was frequently inactivated in multiple tumor types. In addition, we wanted to determine if RORA could also be functioning as a stress-response gene similar to other large CFS genes. We found that RORA is expressed in different human tissues and only two out of its four isoforms are actually transcribed. Then, we showed that the level of RORA is downregulated in breast, ovarian and prostate cancer samples, including primary tumor and cancer cell lines. We also demonstrated that the expression of RORA can be activated by different types of stress treatment. Our data from this study suggested that RORA, another large gene located in a CFS region, plays a role in cellular stress response and is involved in human tumorigenesis.

Results

RORA is derived from the middle of the FRA15A (15q22.2) CFS

RORA is derived from human chromosomal band 15q22.2, which also contains the FRA15A CFS. In order to determine if RORA is located with the FRA15A CFS region, we selected a BAC clone (CTD-2034M3) that spans the approximate center of RORA. This BAC was labeled with biotin, and hybridized to metaphase preparations produced from lymphocytes cultured in the presence of 0.4 μM APC for 24 h. In 20 clear metaphases (out of over 700 analysed) with good discernible breakage/decondensation within FRA15A, BAC CTD-2034M3 was found to localize proximal to the region of B/D in 12 metaphases and distal to B/D in another 8. Figure 1 show representative hybridizations demonstrating that RORA lies within the FRA15A region. Based upon the number of times that this BAC hybridized proximal as compared to distal to the region of B/D, we can determine that the center of the RORA gene lies close to the center of the FRA15A CFS. We also determined that the FRA15A has a relatively low level of expression as we had to analyze over 700 metaphases in order to find sufficient cells with B/D at 15q22.2.

RORA and its specific isoforms are expressed in human normal tissues

We first investigated whether RORA was expressed in normal human tissues, including the brain, breast, liver, ovary and prostate, and if so, which of the four RORA isoforms was present. Using semiquantitative reverse

transcription-polymerase chain reaction (RT–PCR) and the universal primers for the RORA gene, which amplify one 250 bp fragment located from +750 to +1005 of the RORA transcript, it was shown that the human RORA gene was transcribed in all these tissue samples (Figure 2). Using specific primers for each of the four different RORA isoforms, we demonstrated that only RORA1 and RORA4 were transcriptionally expressed in normal tissues (Figure 2). On the other hand, the level of RORA2 and RORA3 was undetectable, even if the amount of cDNA template or the number of PCR cycles was increased. The same expression pattern of RORA was also found in the normal breast epithelial cell line MCF12F (data not shown). These results suggest that RORA is expressed in normal human tissues and RORA1 and RORA4 are the main transcripts of the RORA gene. These are consistent with the results obtained from other research groups (Lau *et al*., 1999; Chauvet *et al*., 2002; Migita *et al*., 2004).

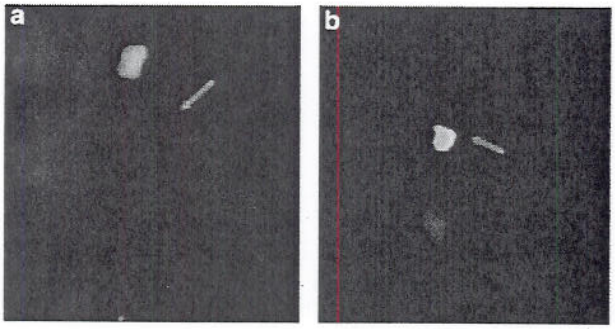


Figure 1 Depiction of FISH results obtained with a BAC clone crossing the middle of RORA and determined to be crossing FRA15A. BAC clone CTD-2034M3 was labeled with biotin and hybridized to normal human lymphocytes treated for 24 h with 0.4 μM APC. Twenty metaphases with clear breakage/decondensation at 15q22.2 were scored. The hybridization signal appeared proximal to the break in 12 metaphases and distal in eight, showing that RORA is located in the approximate center of FRA15A. (**a**) Representative metaphase with the hybridization signal appearing distal to the break. (**b**) Representative metaphase with the hybridization signal appearing proximal to the break.

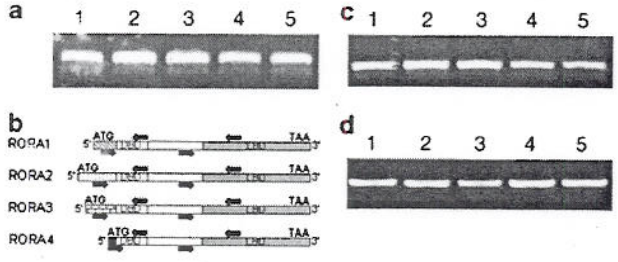


Figure 2 Transcriptional level of RORA and its isoforms in various normal tissues. Lane 1, brain; lane 2, breast; lane 3, liver; lane 4, ovary; and lane 5, prostate. Total RNA was prepared from normal human tissues and cDNA was generated. Semiquantitative RT-PCR was performed using the universal primers for all RORA isoforms and the specific primers for each isoforms to measure the level of RORA. (**a**) RORA universal primers, (**b**) A schematic diagram showing the four different isoforms of RORA. The bold arrows show the positions of specific primers for each isoform and the universal primers for all isoforms. DBD, DNA-binding domain; LBD, ligand-binding domain. (**c**) RORA1 (isoform 1) primers and (**d**) RORA4 (isoform 4) primers.

RORA is downregulated in cancer, including the breast, ovary and prostate cancer samples

In order to understand the role of RORA in the process of human tumorigenesis, the expression level of RORA was measured in several different types of human cancer samples, either in primary tumors or in tumor-derived cell lines. Total RNA was isolated from the breast, ovary and prostate samples. cDNA was prepared and the transcriptional level of RORA was examined with semiquantitative RT–PCR using universal RORA primers designed to be able to amplify all four RORA transcripts, if there is any. The results showed that in most cancer samples, the transcriptional level of RORA was significantly decreased compared to that in normal control samples (Figure 3a–c). To confirm that the downregulation of RORA is specific and not because of the bystander effect of an unstable genomic structure, the expression of another two genes, FOXB1 and NARG2, which are located in the same fragile site as RORA, was examined in these cancer samples. It was demonstrated that, in these cancer-derived samples, the transcriptional level of FOXB1 and NARG2 was not affected compared to the level in controls (Figure 3d).

The expression of RORA is activated by aphidicolin treatment

As the human RORA gene is located in the middle of the CFS FRA15A, it would be interesting to know if the expression status of RORA is affected when the FRA15A CFS was induced. It was well known that the expression of CFS can be induced by aphidicolin

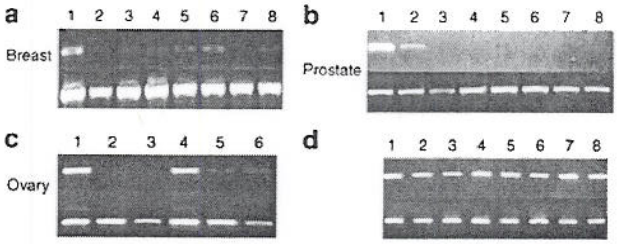


Figure 3 RORA is downregulated in different types of cancer cell lines and human primary cancer samples. Total RNA was extracted and reverse transcribed into cDNA. PCR was performed with the universal primers to examine the transcriptional level of RORA. (**a**) Breast cancer cell lines: lane 1, MCF12F; lane 2, MCF7; lane 3, MDA157; lane 4, UACC893; lane 5, ZR75; lane 6, MDA435; lane 7, T47D; and lane 8, BT474. Top row RORA, and bottom row actin. (**b**) Prostate cancer cell lines and primary tumor samples: lane 1, normal prostate control; lane 2, DU145; lane 3, PC3; lane 4, LNCaP; and Lanes 5–8, primary prostate tumor tissues. Top row RORA, and bottom row actin. (**c**) Ovary cancer cell lines, Lane 1, normal ovarian epithelium control (OSE); lane 2, OV167; lane 3, OV177; lane 4, OV202; lane 5, OVCAR5; and lane 6, SKOV3. Top row RORA; and bottom row actin. (**d**) Breast; lane 1, MCF12F; lane 2, MCF7; lane 3, MDA157; lane 4, UACC893; lane 5, ZR75; lane 6, MDA435; lane 7, T47D; and Lane 8, BT474. Top row FOXB1, and Bottom row NAGR2.

(APC), an inhibitor of DNA polymerase alpha and delta (Glover *et al.*, 1984). The transcriptional level of RORA was checked in the normal breast epithelial cell MCF12F treated with 0.2, 0.4 and 0.8 μM APC for 24 h. It was observed that a step-wise increase in RORA mRNA with the highest level of RORA expression achieved at 0.8 μM APC treatment (Figure 4). On the other hand, the transcriptional level of FOXB1, which is adjacent to RORA on Chromosome 15, was stable under these APC stress conditions (Figure 4). Meanwhile, the abundance of actin mRNA was also not significantly affected. These results suggest that RORA is specifically upregulated when the expression of CFS increases.

RORA also responds to other types of stress

In order to determine whether the activation effect of cellular stresses on RORA gene expression is a general phenomenon, the effect of different types of stresses, including UV, H_2O_2 and methyl methanesulfonate (MMS), on the amount of RORA transcript and protein in MCF12F cells was observed. Three different dosages of each stress agent were used, specifically 10, 20 and 50 J/M^2 for UV; 100, 200 and 500 μM for H_2O_2; and 0.001, 0.005 and 0.01% for MMS. MCF12F cells were treated with each stress condition for 24 h before cellular total RNA and protein were collected. Using semi-quantitative RT–PCR and Western blotting assay, it was shown that there were significant increases of RORA transcript and protein under the stress conditions compared to untreated cells. The results from either semiquantitative RT–PCR or Western blotting assay were consistent. Meanwhile, the expression level of both FOXB1 and NARG2 genes was not affected by any of these stress conditions (data not shown). This effect was specific, since the level of actin transcript and protein was not affected by the treatments (Figure 5).

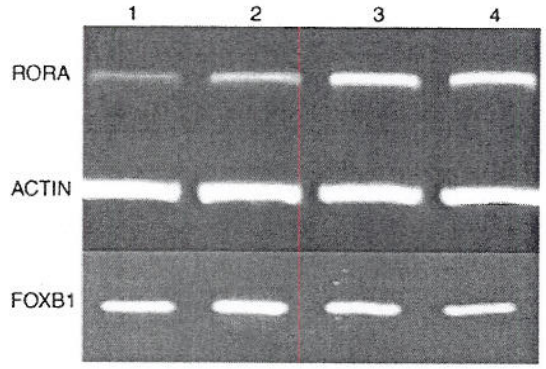


Figure 4 Effect of APC on the transcription of RORA in MCF12F cells. MCF12F cells were treated with various doses of APC for 24 h before total RNA was extracted and cDNA was prepared. PCR was set up to check the transcriptional level of RORA, FOXB1 and actin, using the universal primers for RORA, primers for FOXB1 and the control primers for actin. Lane 1, cell without APC treatment; lane 2, with APC 0.2 μM; lane 3, with APC 0.4 μM; and lane 4, with APC 0.8 μM.

Overexpression of RORA inhibits the cellular growth

In order to elucidate the biological function of RORA in the development of human cancers, the RORA gene was re-introduced into MCF12F cells, and its effect on the cell proliferation was observed. As shown in Figure 6, MCF12F cells were transfected with pcDNA3-RORA4, as RORA4 is one of the predominant isoforms of

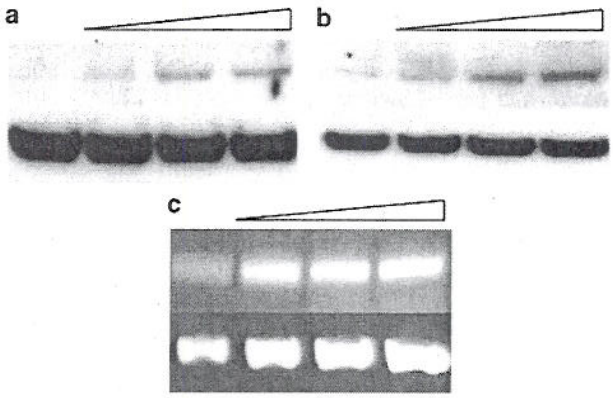


Figure 5 Expression of RORA in MCF12F cells is activated by different types of stress treatments. (**a**) The effect of UV on the protein level of RORA. MCF12F cells were treated with UV at 10, 20 and 50 J/m^2, and total protein was prepared after another 24 h incubation. (**b**) The effect of MMS on the protein level of RORA. MCF12F cells were treated with MMS at 0.001, 0.005 and 0.01% for 24 h before the total cellular protein was extracted. The level of RORA was examined with anti-RORA antibody. (**c**) The effect of H_2O_2 on the transcriptional level of RORA. MCF12F cells were treated with H_2O_2 at 100, 200 and 500 μM for 24 h before the total RNA was extracted and cDNA was prepared. PCR was preformed using the universal primers for RORA to detect the level of RORA.

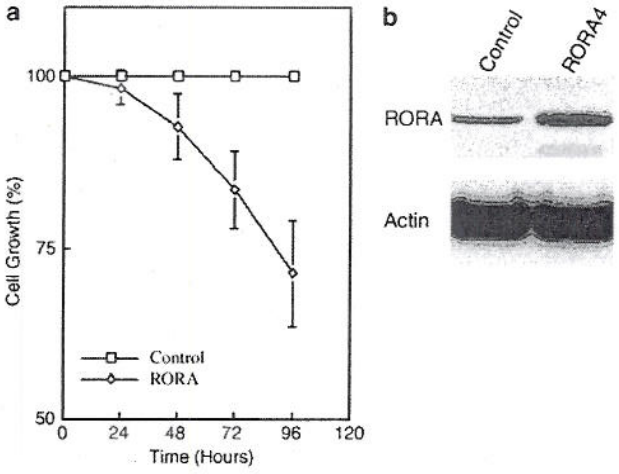


Figure 6 Effect of RORA overexpression on cellular growth. (**a**) MCF12F cells were plated and incubated overnight before transfection. The plasmids (pcDNA3 as control and pcDNA3-RORA4) were transfected into cells using a Lipofectamine 2000 transfection kit following the manufacturer's protocol. The cell number was counted 24, 48, 72 and 96 h later. All results are the average of at least three independent experiments with s.d. shown by bars. (**b**) The level of RORA in pcDNA3-RORA4 transfectants detected by the Western blotting assay. Top row RORA; and bottom row actin.

RORA expressed (Chauvet *et al.*, 2002) and RORA1 is more specifically produced in the central nervous system (Matysiak-Scholze and Nehls, 1997). The cell number was counted 24, 48, 72 and 96 h afterwards. Compared to the growth of vector-expressing clones, the RORA-expressing clones proliferated more slowly. Using the number of vector-expressing clones as the control, the growth of RORA expression clones was 98.4, 92.7, 83.6 and 71.4%, at 24, 48, 72 and 96 h, respectively after transfection (Figure 6). This suggested that increase of RORA expression affected the cell proliferation.

Discussion

CFSs are highly unstable genomic regions that are apparently present in all individuals. While they are characterized utilizing an *in vitro* assay of chromosomal decondensation/breakage induced by inhibitors of DNA replication, their apparent *in vivo* significance is that they predispose chromosomes to breakage and rearrangement, especially in developing cancer cells (Huebner *et al.*, 1998; Smith *et al.*, 1998). The four most active of the CFS regions are FRA3B (3p14.2), FRA16D (16q23.2), FRAXB (Xp22.31) and FRA6E (6q26). Three of these regions, FRA3B, FRA16D and FRA6E, are consistently deleted during the development of many different cancers. Spanning the most unstable regions within each of these three CFS regions are genes that themselves span very large genomic regions, FHIT, WWOX and Parkin, respectively. There are frequent deletions and other alterations in each of these genes in multiple cancers and the proteins encoded by these genes are frequently not expressed in these same cancers (Buttel *et al.*, 2004). However, these genes are not traditional mutational targets in cancer as there are very few cancers that have point mutations in these genes. This may be because their chromosomal locations within the highly unstable CFSs make larger deletions and other alterations predominant. Although FHIT is not a traditional tumor suppressor gene, it has been demonstrated that the inactivation of this gene results in cells being much more tumor prone, and this can be reversed by putting FHIT back into FHIT−/− cells (Gopalakrishnan *et al.*, 2003). The re-introduction of WWOX and Parkin into cell lines that do not express them can sometimes inhibit the growth of these cells as well as result in greater sensitivity to apoptotic induction (Bednarek *et al.*, 2001; Jiang *et al.*, 2004; Wang *et al.*, 2004).

As several of the CFS regions are associated with extremely large genes, we have begun to systematically examine other large genes that are derived from regions known to contain a CFS to determine if they are also CFS genes. This strategy enabled us to identify the 730 kb RORA gene as another gene that spans a CFS region. The RORA gene contains 11 exons that together make a 1.5 kb final processed transcript (Jetten and Ueda, 2002). The cytogenetic results from this study demonstrate that RORA is derived from the approximate middle of the FRA15A (15q22.2) CFS. In contrast to FRA3B, FRA16D and FRA6E, which are highly active frequently expressed CFSs, FRA15A is expressed at very low frequencies. We therefore did not characterize the entire FRA15A region, but have merely delineated the position of the center of this CFS region.

We have shown that RORA is universally expressed in various normal tissues, including the brain, breast, liver, ovary and prostate. Human RORA produces four isoforms that are identical, except for their amino-terminus (Giguere *et al.*, 1994; Hamilton *et al.*, 1996). These isoforms are generated by a combination of alternative promoter usage and exon splicing. The RORA1 isoform was reported to be specifically produced in the central nervous system (Matysiak-Scholze and Nehls, 1997), but in this study, we found that both RORA1 and RORA4 were present in an organ other than the brain. This is consistent with the results reported previously. The presence of both RORA1 and RORA4 in various organs suggested that they might be critical in the regulation of certain physiological processes by controlling the expression of their target genes.

It has been reported that RORA appears to be crucial for many cellular physiological processes that occur in tissues such as the cerebellum, adipose, muscle and bone (Jetten, 2004). However, the precise role that RORA plays *in vivo* remains to be elucidated, and as a member of the nuclear receptor family, its cognate ligand remains to be identified (Giguere, 1999). In the current study, we found that the transcriptional level of RORA was decreased in several different types of cancer, including cancer-derived cell lines and primary tumors. We have now demonstrated that the expression of RORA was activated when the cells were exposed to various types of stress, including UV, MMS and H_2O_2. This suggests that RORA could contribute to the development of cancer, and owing to its inactivation, the normal stress response system that RORA is a part of is no longer functional.

The altered level of RORA might affect the expression of its downstream target genes that are normally used to execute its biological function(s). It was previously demonstrated that RORA binds as a monomer or homodimer to regulatory ROREs in the promoter region of its target genes, which include one core motif AGGTCA or two direct AGGTCA repeats spaced by two nucleotides preceded by a six nucleotide long AT-rich sequence (Giguere *et al.*, 1994). Several genes have been identified as potential target genes of RORA including the CDK inhibitor p21 and N-myc (Dussault and Giguere, 1997), which are closely related to human cancer development. In this study, a significantly inhibited cell growth was observed when the RORA gene was re-introduced into MCF12F cells. It is suggested that the level of RORA expression does affect the cell proliferation. Further experiments are required to understand how RORA itself (and its downstream target genes) is involved in the regulation of the cell growth. It has been suggested that RORA significantly decreases cell proliferation and affects cell

cycle progression through the modulation of cell cycle-related genes in DU145 androgen-independent prostate cancer cells (Moretti *et al.*, 2001).

In a previous study, it was demonstrated that hypoxia increased the amount of RORA transcripts (Chauvet *et al.*, 2004). As hypoxia is an important component of many physiological and pathological processes, this suggests that RORA may be involved in the cellular stress response network. It has been well documented that the response to multiple types of genetic damage is mediated by the DNA damage checkpoint pathway (Nyberg *et al.*, 2002; Laiho and Latonen, 2003). Recently, another study showed that the components of the DNA damage checkpoint pathway controlled the expression of two CFS genes FHIT and WWOX in response to UV treatment (Ishii *et al.*, 2005). It was also reported that the expression of CFSs themselves was affected by the S phase and G2/M cellular DNA damage checkpoint protein ATR (ataxia-telangiectasia and Rad3-related; Casper *et al.*, 2002). This mechanism can be activated when cells are treated with UV or DNA replication blocking agents, such as APC. The deficiency of ATR, but not ATM (ataxia-telangiectasia mutated), caused a significant increase of chromosomal instability after APC treatment. Both the altered expression of CFS genes and increased fragility would result in increased susceptibility to cancer. Further experiments will need to be performed in order to define the connection between RORA and certain checkpoint component(s) and between RORA and its downstream target genes.

RORA mRNA is widely expressed. This contrasts with the more selective expression of RORB and RORC, two other members of the same family (Jetten, 2004). It was reported that RORB is expressed specifically in the brain (Carlberg *et al.*, 1994), and RORC is found at high levels in the skeletal muscle (Hirose *et al.*, 1994). However, we do not have any evidence to rule out the possibility that RORB and RORC also play some role in tumorigenesis *in vivo*. Actually, it was shown that exogenous expression of either of the two isforms of RORC in T-cell hybridomas inhibited interleukin-2 and Fas ligand expression and blocked T-cell receptor-induced cell proliferation and apoptosis (He *et al.*, 1998; Littman *et al.*, 1999). Recently, it was reported that RORC is a common integration site in type B leukemogenic virus-induced T-cell lymphomas (Broussard *et al.*, 2004).

In summary, we have now shown that the very large RORA gene is actually derived from within the most active portion of the FRA15A CFS and that this gene is frequently inactivated in cancers that arise from different types of human tissues. RORA was previously shown to be involved in cellular response to hypoxia, and here we demonstrated changes in the amount of RORA gene product in cells exposed to a variety of different cellular stresses, including UV, MMS and H_2O_2. Thus, RORA is another very large CFS gene that is inactivated in multiple tumors. In addition, RORA appears to play a critical role in responses to cellular stress, lending further support to the idea that the large CFS genes function as part of a highly conserved stress response network that is uniquely susceptible to genomic instability in cancer cells.

Materials and methods

Clone selection

A bacterial artificial chromosome (BAC) clone was selected based on its localization within the RORA gene according to the UCSC Human Genome Database. The selected clone for RORA (CTD-2034M3) spans the approximate center of the gene. BAC clone CTD-2034M3 was obtained from Invitrogen and grown according to published procedures. The BAC clone was verified by PCR with primers derived from markers that were spanned by the BAC clone. DNA was then isolated from individual colonies that amplified the correct sized fragment using PCR as described previously.

Cell culture and transfection

The normal breast epithelium cell line MCF12F was obtained from American Type Culture Collection (Rockville, MD, USA). MCF12F cells were routinely grown in DMEM/Ham's F12 supplemented with 20 ng/ml epidermal growth factor (EGF), 100 ng/ml cholera toxin, 0.01 mg/ml insulin, 500 ng/ml hydrocortisone and 5% horse serum. All the other cell lines were maintained in DMEM medium supplemented with 10% fetal calf serum (FCS), 2 mM L-glutamine, and 100 U/ml penicillin, 100 μg/ml streptomycin at 37°C in a humidified atmosphere of 5% CO_2. Metaphase cell preparations were prepared from mitogen-stimulated peripheral blood cultures obtained from normal individuals. Cultures were established with 9.5 ml RPMI 1640, 10% fetal bovine serum, 100 U/ml penicillin, 100 μg/ml streptomycin, 0.5 ml lymphocyte-rich blood, and 10 μg/ml PHA (Irvine Scientific, Santa Ana, CA, USA) and incubated at 37°C in 5% CO_2 for 72 h. Twenty-four hours prior to harvest, select cultures were inoculated with 0.4 μM APC (Sigma, St Louis, MO, USA). Cell harvest and slide preparation followed routine cytogenetic techniques. The transient transfection was performed using Lipofectamine 2000 (Invitrogen, Carlsbad, CA, USA) according to the manufacturer's protocol. The construct pcDNA3-RORA4, which overexpresses the product of RORA4, was generously provided by Dr V Giguere at McGill University Health Center.

Cellular stress treatment

For APC and MMS (Sigma, St Louis, MO, USA) treatment, 2×10^5 MCF12F cells were plated in a 60 mm Petri dish and incubated at 37°C for overnight. Next day, the fresh medium was exchanged and the chemical was directly added into the medium with different concentrations. Cells were harvested after 24 h incubation and total RNA/protein was prepared. For UV treatment, 2×10^5 MCF12F cells were plated in a 60 mm Petri dish. After overnight incubation at 37°C, the medium was extracted and the cells were exposed to 254 nm UV light at different intensities in a UV crosslinker (Fisher model FB-UVXL-1000 at ~2400 μW/cm^2). Fresh medium was added and cells were incubated for another 24 h before the RNA or protein extraction.

Fluorescence in situ *hybridization*

Fluorescence *in situ* hybridization (FISH) can be preformed to detect the expression of CFS. RORA spans ~730 kb within chromosomal band 15q22.2. A single BAC clone from within

the approximate center of this gene was selected to test whether this gene was derived from within the CFS that has been localized to the same chromosomal band. Purified BAC DNA was biotin labeled using a BioNick translation kit (Invitrogen) according to the manufacturer's protocols. The probe was precipitated and hybridized to APC-treated metaphase chromosomes. After overnight hybridization at 37°C, slides were washed twice in 2 × SSC (pH 7.0), twice in 50% 2 × SSC/50% formamide and twice in 2 × SSC (pH 7.0). All solutions were at 40°C and each wash was 5 min. Slides were then incubated at room temperature in 75 μl of 5% bovine serum albumin/4 × SSC + 0.2% Tween-20 solution for 5 min. The detection of the probe signal was performed by applying 75 μl of an avidin/5% bovine serum albumin mixture to each slide, incubating at 37°C for 15 min and washing in four changes of 4 × SSC + 0.2% Tween-20 solution at 40°C for 3 min each. Signals were amplified by applying 75 μl of 5% normal goat serum/4 × SSC + 0.2% Tween-20 solution to each slide and incubating for 5 min at room temperature. An anti-avidin/5% normal goat serum solution (75 ml) was then applied to each slide and then incubated at 37°C for 15 min, and washed in four changes of 4 × SSC + 0.2% Tween-20 solution at 40°C for 3 min each. Chromosomes were counterstained with DAPI (Vector Laboratories, Burlingame, CA, USA). Photomicroscopy was performed using a Zeiss Axioplan flurorescence microscope equipped with MacProbe software (Applied Imaging, San Jose, CA, USA).

Semiquantitative RT–PCR

Total RNA was isolated from cell lines and primary tumor samples with a Versagene™ RNA Purification Kit (Gentra, Minneapolis, MN, USA) according to the manufacturer's protocol. Reverse transcription was performed at 50°C for 1 h in a total volume of 20 μl using a ThemoScript™ RT-PCR Kit (Invitrogen) according to the manufacturer's protocol. The RORA isoform-specific PCR primer sequences were deduced from the published sequences of the human and mouse RORA1 and RORA4, and the human RORA2 and RORA3 cDNA (Genebank Accession numbers U04897, U53228, L14611, Y08640, U04898 and U04899). The universal primers used for all RORA transcripts are RORA-5′, 5′-GTCACGCAGCTTCTACCTGGAC-3′ and RORA-3′, 5′-GTGTTGTTCTGAGAGTGAAAGGCAGG-3′. Each of the RORA1-4 transcripts was analysed using the RORA-3′ isoprimer, 5′-AACAGTTCTTCTGACGAGGACAGG-3′ for all the isoforms, and the RORA1-5′, 5′-GAGGTATCTCAGTCACGAAC-3′ (183 bp product) for RORA1; the RORA2-5′ primer, 5′-CAGTGTATCCTGTCTTCAGG-3′ (274 bp product) for RORA2; the RORA3-5′ primer, 5′-ACATAAACTGGGATGGAGCC-3′ (234 bp product) for RORA3; and the RORA4-5′ primer, 5′-TGTGATCGCAGCGATGAAAG-3′ (170 bp product) for RORA4. The primers for FOXB1 are 5′-CCGCCCTACTCGTACATCTC-3′ and 5′-CGGGATCTTGATGAAGCAGT-3′. The primers for NARG2 are 5′-CCCTTGAAGTTTGAGGAGGA-3′ and 5′-GTCGCAACAGACTGGCAATA-3′. The amount of actin transcripts was used as the loading control, which was detected with the control primers actin-5′, 5′-ATGAGGTAGTCAGTCAGGTC-3′; and actin-3′, 5′-GCTCCGGCATGTGCAAGG-3′. The thermal cycle conditions included one cycle at 95°C for 10 min, 30 cycle of (95°C for 30 s, 55°C for 30 s, 72°C for 30 s) and one cycle at 72°C for 10 min.

Western blotting assay

Cultured cells were harvested and incubated in lysis buffer (25 mM Tris-phosphate, 2 mM DTT, 2 mM diaminocyclohexane tetraacetic acid, 10% glycerol, 1% Triton X-100, 5 mM PMSF) for 10 min on ice. After centrifugation for 10 min at 13 000 r.p.m., the supernatant was collected and the protein concentration was determined using a spectrophotometer. Samples containing 40 μg of total protein were electrophoresed on 10% SDS–PAGE gels, and transferred to PVDF membranes. Anti-RORA antibody was obtained from Santa Cruz Biotechnology (Lot no. sc-6062) and anti-actin antibody was from NeoMakers (Lot no. 1295P405H). The secondary antibody was used to visualize the protein expression by an enhanced chemiluminescence detection kit (Amersham Biosciences, Buckinghamshire, UK; Catalog no. 25-0062-62) according to the manufacturer's instructions.

References

Arlt MF, Casper AM, Glover TW. (2003). *Cytogenet Genome Res* **100**: 92–100.

Becker NA, Thorland EC, Denison SR, Phillips LA, Smith DI. (2002). *Oncogene* **21**: 8713–8722.

Becker-Andre M, Andre E, DeLamarter JF. (1993). *Biochem Biophys Res Commun* **194**: 1371–1379.

Bednarek AK, Keck-Waggoner CL, Daniel RL, Laflin KJ, Bergsagel PL, Kiguchi K *et al.* (2001). *Cancer Res* **61**: 8068–8073.

Broussard DR, Lozano MM, Dudley JP. (2004). *J Virol* **78**: 4943–4946.

Buttel I, Fechter A, Schwab M. (2004). *Ann NY Acad Sci* **1028**: 14–27.

Carlberg C, Hooft van Huijsduijnen R, Staple JK, DeLamarter JF, Becker-Andre M. (1994). *Mol Endocrinol* **8**: 757–770.

Casper AM, Nghiem P, Arlt MF, Glover TW. (2002). *Cell* **111**: 779–789.

Chang NS, Doherty J, Ensign A, Lewis J, Heath J, Schultz L *et al.* (2003). *Biochem Pharmacol* **66**: 1347–1354.

Chauvet C, Bois-Joyeux B, Berra E, Pouyssegur J, Danan JL. (2004). *Biochem J* **384**: 79–85.

Chauvet C, Bois-Joyeux B, Danan JL. (2002). *Biochem J* **364**: 449–456.

Druck T, Hadaczek P, Fu TB, Ohta M, Siprashvili Z, Baffa R *et al.* (1997). *Cancer Res* **57**: 504–512.

Dumon KR, Ishii H, Fong LY, Zanesi N, Fidanza V, Mancini R *et al.* (2001). *Proc Natl Acad Sci USA* **98**: 3346–3351.

Dussault I, Giguere V. (1997). *Mol Cell Biol* **17**: 1860–1867.

Finnis M, Dayan S, Hobson L, Chenevix-Trench G, Friend K, Ried K *et al.* (2005). *Hum Mol Genet* **14**: 1341–1349.

Gemma A, Hagiwara K, Ke Y, Burke LM, Khan MA, Nagashima M *et al.* (1997). *Cancer Res* **57**: 1435–1437.

Giguere V. (1999). *Endocr Rev* **20**: 689–725.

Giguere V, Tini M, Flock G, Ong E, Evans RM, Otulakowski G. (1994). *Genes Dev* **8**: 538–553.

Glover TW, Berger C, Coyle J, Echo B. (1984). *Hum Genet* **67**: 136–142.

Gopalakrishnan VK, Banerjee AG, Vishwanatha JK. (2003). *Pancreatology* **3**: 293–302.

Hamilton BA, Frankel WN, Kerrebrock AW, Hawkins TL, FitzHugh W, Kusumi K *et al.* (1996). *Nature* **379**: 736–739.

He YW, Deftos ML, Ojala EW, Bevan MJ. (1998). *Immunity* **9**: 797–806.

Hirose T, Smith RJ, Jetten AM. (1994). *Biochem Biophys Res Commun* **205**: 1976–1983.

Huebner K, Garrison PN, Barnes LD, Croce CM. (1998). *Annu Rev Genet* **32**: 7–31.
Huebner K, Hadaczek P, Siprashvili Z, Druck T, Croce CM. (1997). *Biochim Biophys Acta* **1332**: M65–M70.
Ishii H, Mimori K, Inageta T, Murakumo Y, Vecchione A, Mori M *et al.* (2005). *Mol Cancer Res* **3**: 130–138.
Jetten AM. (2004). *Curr Drug Targets Inflamm Allergy* **3**: 395–412.
Jetten AM, Ueda E. (2002). *Cell Death Differ* **9**: 1167–1171.
Jiang H, Ren Y, Zhao J, Feng J. (2004). *Hum Mol Genet* **13**: 1745–1754.
Krummel KA, Denison SR, Calhoun E, Phillips LA, Smith DI. (2002). *Genes Chromosomes Cancer* **34**: 154–167.
Laiho M, Latonen L. (2003). *Ann Med* **35**: 391–397.
Lau P, Bailey P, Dowhan DH, Muscat GE. (1999). *Nucleic Acids Res* **27**: 411–420.
Lee WH, Murphree AL, Benedict WF. (1984). *Nature* **309**: 458–460.
Littman DR, Sun Z, Unutmaz D, Sunshine MJ, Petrie HT, Zou YR. (1999). *Cold Spring Harb Symp Quant Biol* **64**: 373–381.
Ludes-Meyers JH, Bednarek AK, Popescu NC, Bedford M, Aldaz CM. (2003). *Cytogenet Genome Res* **100**: 101–110.
Matsuyama A, Shiraishi T, Trapasso F, Kuroki T, Alder H, Mori M *et al.* (2003). *Proc Natl Acad Sci USA* **100**: 14988–14993.
Matysiak-Scholze U, Nehls M. (1997). *Genomics* **43**: 78–84.
Migita H, Satozawa N, Lin JH, Morser J, Kawai K. (2004). *FEBS Lett* **557**: 269–274.
Missbach M, Jagher B, Sigg I, Nayeri S, Carlberg C, Wiesenberg I. (1996). *J Biol Chem* **271**: 13515–13522.
Moretti RM, Marelli MM, Motta M, Polizzi D, Monestiroli S, Pratesi G *et al.* (2001). *Prostate* **46**: 327–335.
Nau MM, Brooks Jr BJ, Carney DN, Gazdar AF, Battey JF, Sausville EA *et al.* (1986). *Proc Natl Acad Sci USA* **83**: 1092–1096.
Nyberg KA, Michelson RJ, Putnam CW, Weinert TA. (2002). *Annu Rev Genet* **36**: 617–656.
Pawlikowski M, Kunert-Radek J, Winczyk K, Melen-Mucha G, Gruszka A, Karasek M. (1999). *Adv Exp Med Biol* **460**: 369–372.
Schrader M, Danielsson C, Wiesenberg I, Carlberg C. (1996). *J Biol Chem* **271**: 19732–19736.
Smith DI, Huang H, Wang L. (1998). *Int J Oncol* **12**: 187–196.
Sutherland GR, Baker E, Richards RI. (1998). *Trends Genet* **14**: 501–506.
Sutherland GR, Richards RI. (1995). *Curr Opin Genet Dev* **5**: 323–327.
Wang F, Denison S, Lai JP, Philips LA, Montoya D, Kock N *et al.* (2004). *Genes Chromosomes Cancer* **40**: 85–96.